MNRAS **454,** 1266–1289 (2015) doi:10.1093/mnras/stv2042

Towards optimal estimation of the galaxy power spectrum

Robert E. Smith[1,2★] and Laura Marian[1★]

[1]*Department of Physics and Astronomy, University of Sussex, Brighton BN1 9QH, UK*
[2]*Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str.1, Postfach 1523, D-85740 Garching, Germany*

Accepted 2015 August 28. Received 2015 August 28; in original form 2015 March 30

ABSTRACT

The galaxy power spectrum encodes a wealth of information about cosmology and the matter fluctuations. Its unbiased and optimal estimation is therefore of great importance. In this paper, we generalize the framework of Feldman et al. (1994) to take into account the fact that galaxies are not simply a Poisson sampling of the underlying dark matter distribution. Besides finite survey-volume effects and flux limits, our optimal estimation scheme incorporates several of the key tenets of galaxy formation: galaxies form and reside exclusively in dark matter haloes; a given dark matter halo may host several galaxies of various luminosities; galaxies inherit part of their large-scale bias from their host halo. Under these broad assumptions, we prove that the optimal weights *do not* explicitly depend on galaxy luminosity, other than through defining the maximum survey volume and effective galaxy density at a given position. Instead, they depend on the bias associated with the host halo; the first and second factorial moments of the halo occupation distribution; a selection function, which gives the fraction of galaxies that can be observed in a halo of mass M at position $\boldsymbol{r}$ in the survey; and an effective number density of galaxies. If one wishes to reconstruct the matter power spectrum, then, provided the model is correct, this scheme provides the only unbiased estimator. The practical challenges with implementing this approach are also discussed.

Key words: large-scale structure of Universe.

1 INTRODUCTION

The power spectrum of matter fluctuations, or equivalently the two-point correlation function, is a fundamental tool for constraining the cosmological parameters. It contains detailed information about the large-scale geometrical structure of space–time, the constituents of energy-density and their evolution with redshift, and also provides us with information about the primordial scalar fluctuation spectrum. However, we do not directly observe the matter density field; instead, we observe galaxy angular positions and measure radial velocities, or redshifts, from spectra. Given a galaxy redshift survey, two things are crucial: how to obtain an unbiased estimate of the information in the matter fluctuations; and obtaining an estimate that has the highest signal to noise possible, i.e. an optimal measurement.

The first point may be rephrased as the need to understand the relation between galaxy and matter fluctuations – more commonly referred to as galaxy bias. The second point, that of optimality, in fact also relies on our understanding of bias, since only through knowing how the galaxies are embedded in the mass distribution can one devise efficient survey strategies; for example, if all galaxies formed in pairs then one would only require information about one galaxy from each pair to obtain all of the useful cosmological information.

The development of galaxy correlation functions as a tool for constraining the cosmological model was first realized by Peebles and collaborators in a series of pioneering papers in the 1970s (Hauser & Peebles 1973; Peebles 1973, 1974, 1975; Peebles & Hauser 1974; Peebles & Groth 1975; Groth & Peebles 1977; Seldner & Peebles 1977, 1978, 1979; Fry & Peebles 1978, 1980). Subsequent studies built on this, and estimators were developed that took better account of fluctuations in the mean number density of galaxies (Davis & Peebles 1983; Hamilton 1993; Landy & Szalay 1993; Bernstein 1994). These new estimators were only optimal in the case that the clustering was very weak and when galaxies represented a Poisson sampling of the underlying matter fluctuations.

The development of techniques for the direct estimation of the galaxy power spectrum began in earnest in the early 1990s (Baumgart & Fry 1991; Peacock & Nicholson 1991; Fisher et al. 1993). This culminated in the seminal work of Feldman et al. (1994, hereafter FKP). In their seminal approach, galaxies were assumed to be a Poisson sampling of the mass density field. They showed that provided one subtracted

★ E-mail: r.e.smith@sussex.ac.uk (RES); l.marian@sussex.ac.uk (LM)

an appropriate shot-noise term, and deconvolved for the survey window function, one could obtain an unbiased estimate of the matter power spectrum. Subsequent analysis focused on obtaining quadratic and decorrelated band estimates (Vogeley & Szalay 1996; Hamilton 1997a,b, 2000; Hamilton & Tegmark 2000).

In the last two decades, our understanding of galaxy formation has made rapid progress and the current best models strongly suggest that galaxies are not related to the underlying dark matter in the simple way that was envisioned in FKP (White & Rees 1978; White & Frenk 1991; Kauffmann et al. 1999; Benson et al. 2000; Springel et al. 2005). Furthermore, improved observational studies have subsequently discovered that galaxy clustering is in fact dependent on a number of properties of the galaxy distribution: e.g. luminosity (Park et al. 1994; Norberg et al. 2001, 2002; Zehavi et al. 2002, 2005, 2011; Swanson et al. 2008), colour (Brown, Webster & Boyle 2000; Zehavi et al. 2002, 2005, 2011; Swanson et al. 2008), morphology (Davis & Geller 1976; Guzzo et al. 1997; Norberg et al. 2002) and stellar mass (Li et al. 2006) etc.

Percival, Verde & Peacock (2004, hereafter PVP) attempted to correct the FKP framework to take into account the effects of luminosity dependent bias. To this end, PVP assumed that the probability of finding a galaxy of a given luminosity in a certain patch of space would be a Poisson variate, whose mean was proportional to the local density of dark matter multiplied by a luminosity-dependent bias factor. Their work demonstrated two important facts: first, that an optimal weighting scheme depended sensitively on the assumptions about the bias; and secondly, if their assumptions about the bias were correct, the FKP method was a biased estimator of the matter power spectrum.

In this paper, we argue that the approach of PVP, whilst qualitatively reasonable, is in fact still at odds with our current understanding of galaxy formation and therefore unlikely to be the *true* optimal estimator. The key ideas from galaxy formation and evolution that we would like to build into our estimator are: galaxies only form in dark matter haloes (White & Rees 1978); haloes can host a number of galaxies of various luminosities; the large-scale bias associated with a given galaxy is largely inherited from the bias of the host dark matter halo. In a recent paper (Smith & Marian 2014, hereafter SM14), we generalized the FKP formalism to account for the clustering of galaxy clusters – which turned out to have a similar mathematical structure to the PVP scheme. We now undertake to generalize the FKP formalism to take into account these ideas from galaxy formation. As we will show, these effects will lead us to a new optimal estimator and method for reconstructing the matter power spectrum.

Before moving on, it is worth noting that current state-of-the-art galaxy redshift surveys, such as the Baryon Oscillation Spectroscopic Survey (Anderson et al. 2012, 2014a,b, hereafter BOSS), Galaxy And Mass Assembly (Blake et al. 2013, hereafter GAMA) and WiggleZ (Blake et al. 2011), have all used the FKP power spectrum estimation procedure. Future surveys, such as DESI (Levi et al. 2013), Euclid (Laureijs et al. 2011) and SKA (Blake et al. 2004), will have significantly larger volumes and so unbiased and optimized data analysis will be crucial if we are to obtain the tightest constraints on the cosmological parameters.

The paper is broken down as follows. In Section 2, we describe generic properties of a galaxy redshift survey and present a new theoretical quantity, the halo-galaxy double-delta expansion. We explore its statistical properties. In Section 3, we show how one may obtain unbiased estimates of the matter correlation function and power spectrum. In Section 4, we derive the covariance matrix of the fluctuations in the galaxy power spectrum. In Section 5, we derive the optimal weights. In Section 6, we present a new expression for the Fisher information matrix for optimally weighted galaxy power spectra. In Section 7, we enumerate the steps for a practical implementation of this approach. Finally, in Section 8, we summarize our findings and draw our conclusions.

2 SURVEY SPECIFICATIONS AND THE $\mathcal{F}_{\rm g}$–FIELD

2.1 Preliminaries: a generic galaxy redshift survey

Let us begin by defining our fiducial galaxy survey: suppose that we have observed $N_{\rm g}^{\rm tot}$ galaxies and to the ith galaxy we assign a luminosity L_i, redshift z_i and angular position on the sky $\boldsymbol{\Omega}_i = \boldsymbol{\Omega}(\theta_i, \phi_i)$. If we specify the background Friedman-Lemaitre-Robertson-Walker (FLRW) space–time, then we may convert the redshift into a comoving radial geodesic distance $\chi_i = \chi(z_i)$. A galaxy's comoving position vector may now be expressed as $\boldsymbol{r}_i = \boldsymbol{r}(\chi_i, \boldsymbol{\Omega}_i)$.

The survey mask function depends on both the position and luminosity of galaxies, given an adopted flux limit. In this work, we shall take the angular and radial parts of the survey mask function to be separable, though this assumption does not change our results:

$$\Theta(\boldsymbol{r}|L) = \Theta(\boldsymbol{\Omega})\Theta(\chi|L)\,. \tag{1}$$

Note that if the flux limit is not uniform across the survey then the radial function $\Theta(\chi, L)$ would still be a function of the angular position vector $\boldsymbol{\Omega}$, and the survey mask cannot be separated as in the equation above. The angular part of the mask may be written as:

$$\Theta(\boldsymbol{\Omega}) = \begin{cases} 1\,;\ [\boldsymbol{\Omega} \in \{\boldsymbol{\Omega}_\mu\}] \\ 0\,;\ [\text{otherwise}] \end{cases}, \tag{2}$$

where $\{\boldsymbol{\Omega}_\mu\}$ is the set of angular positions that lie inside the survey area. The radial mask function may be written:

$$\Theta(\chi|L) = \begin{cases} 1\,;\ [\chi \le \chi_{\rm max}(L)] \\ 0\,;\quad [\text{otherwise}] \end{cases}, \tag{3}$$

where $\chi_{\rm max}(L)$ is the maximum distance out to which a galaxy of luminosity L could have been detected.

The survey volume for galaxies with luminosity L is simply the integral of the mask function over all space:

$$V_\mu(L) = \int \Theta(\boldsymbol{r}|L)\mathrm{d}V, \tag{4}$$

where $\mathrm{d}V$ is the comoving volume element at position vector $\boldsymbol{r}$ (for a flat universe $\mathrm{d}V = \mathrm{d}^3r = \chi^2\mathrm{d}\Omega\mathrm{d}\chi$). In what follows, it will be also useful to note that the relation $\chi_{\max}(L)$ may be inverted to obtain the minimum galaxy luminosity that could have been detected at radial position $\chi(z)$ in the survey. We shall write this as

$$\left[L_{\min}(\boldsymbol{r})/h^{-2}\mathrm{L}_\odot\right] = 10^{-\frac{2}{5}(m_{\mathrm{lim}}-25-\mathrm{M}_\odot)}\left[d_{\mathrm{L}}(\boldsymbol{r})/\,h^{-1}\mathrm{Mpc}\right]^{-2}, \tag{5}$$

where m_{lim} is the apparent magnitude limit of the survey, $\mathrm{M}_\odot$ is the absolute magnitude of the sun, h is the dimensionless Hubble parameter and d_{L} is the luminosity distance (for a flat universe $d_{\mathrm{L}}(z) = (1+z)\chi(z)$). Thus for any general function $\mathcal{B}(\chi, L)$, we have the useful integral relations:

$$\int_0^\infty \mathrm{d}L \int_0^\infty \mathrm{d}\chi\,\Theta(\chi|L)\mathcal{B}(\chi, L) = \int_0^\infty \mathrm{d}L \int_0^{\chi_{\max}(L)} \mathrm{d}\chi\,\mathcal{B}(\chi, L) = \int_0^\infty \mathrm{d}\chi \int_{L_{\min}(\chi)}^\infty \mathrm{d}L\mathcal{B}(\chi, L)\,. \tag{6}$$

2.2 The halo-galaxy double-delta expansion

Our understanding of galaxy formation tells us that galaxies form exclusively in dark matter haloes, and that each dark matter halo may host several galaxies of various luminosities. It therefore follows that the large-scale bias associated with any given galaxy is directly proportional to the bias of the host halo. We shall mathematically encode these ideas in our density field as follows: our $N_{\mathrm{g}}^{\mathrm{tot}}$ galaxies are distributed inside N_{h} dark matter haloes. Thus, the ith dark matter halo of mass M_i and centre of mass position $\boldsymbol{x}_i$ will host a number of galaxies that depends on its mass, $N_{\mathrm{g}}(M_i)$. The jth galaxy will have a position vector $\boldsymbol{r}_j$ relative to the centre of the halo and a luminosity L_j.

In order to study the statistical properties of the galaxy field in this scenario, we need to simultaneously account for both the spatial distribution of the haloes, as well as the galaxies inside them. We therefore introduce a new function, dubbed the 'galaxy-halo double-delta expansion'. This is a Dirac delta function expansion over the halo positions and masses, as well as over the positions and luminosities of the galaxies inside the haloes. It is written:

$$n_{\mathrm{g}}(\boldsymbol{r}, L, \boldsymbol{x}, M) = \sum_{i=1}^{N_h} \delta^{\mathrm{D}}(\boldsymbol{x}-\boldsymbol{x}_i)\delta^{\mathrm{D}}(M-M_i) \sum_{j=1}^{N_{\mathrm{g}}(M_i)} \delta^{\mathrm{D}}(\boldsymbol{r}-\boldsymbol{r}_j-\boldsymbol{x}_i)\delta^{\mathrm{D}}(L-L_j)\,, \tag{7}$$

where N_{h} is the total number of host dark matter haloes in the survey volume and $N_{\mathrm{g}}(M_i)$ is the total number of galaxies in the ith halo. In the above function, the order of variables is important: $\boldsymbol{r}$ refers to the spatial vector in the galaxy field, L the luminosity, $\boldsymbol{x}$ the spatial vector in the halo field, and M the halo mass. Note that the units of the above function are inverse squared volume, inverse mass and inverse luminosity.[1]

Next, in analogy with SM14, we define a field $\mathcal{F}_{\mathrm{g}}$, which is related to the overdensity of galaxies. Our survey will be finite and will contain masked regions and an apparent magnitude limit m_{lim}. Hence, the overdensity field of galaxies with magnitudes above some threshold luminosity may be written using equation (7) in the following way:

$$\mathcal{F}_{\mathrm{g}}(\boldsymbol{r}) = \int_0^\infty \mathrm{d}L \int \mathrm{d}^3x \int_0^\infty \mathrm{d}M\,\Theta(\boldsymbol{r}|L)\frac{w(\boldsymbol{r}, L, \boldsymbol{x}, M)}{\sqrt{A}}\left[n_{\mathrm{g}}(\boldsymbol{r}, L, \boldsymbol{x}, M) - \alpha\, n_{\mathrm{s}}(\boldsymbol{r}, L, \boldsymbol{x}, M)\right], \tag{8}$$

where $w(\boldsymbol{r}, L, \boldsymbol{x}, M)$ is a weighting function that we will wish to determine in an optimal way, and A is a normalization parameter that will be chosen later.

The function $n_{\mathrm{s}}(\boldsymbol{r}, L, \boldsymbol{x}, M)$ is the random galaxy-halo double-delta expansion. This immediately leads to an important question: what constitutes a random catalogue? Conventionally, in the FKP approach one would distribute the mock galaxies randomly within the survey volume – preserving the number counts as a function of redshift. However, since we know (or have assumed in this model) that galaxies form only inside dark matter haloes, we do not want to remove this property. Instead it is the dark matter haloes which should be randomly distributed, and not the galaxies. Therefore, the function $n_{\mathrm{s}}(\boldsymbol{r}, L, \boldsymbol{x}, M)$ represents the distribution of galaxies in a mock sample *whose dark matter haloes possess no intrinsic spatial correlations*, and have a number density that is $1/\alpha$ of the true galaxy-halo double-delta field. Note that for this random distribution, while the halo centres are not correlated, the galaxies still follow the density distribution inside each dark matter halo. In addition, the haloes possess a mass spectrum and the galaxy luminosities are conditioned on the halo mass.

Both quantities defined by equations (7) and (8) are of central importance and will be extensively used in this paper. It is therefore worthwhile for us to take some time to understand their meaning and how one should employ them to infer the statistical properties of the galaxy density field. This we do in the following section.

[1] We note that this equation is the more rigorous starting point for all halo model calculations of the galaxy field. However, so far as we are aware it has not been written down before. This in part owes to the fact that the galaxy-clustering expressions could be deduced by analogy with the mass clustering. However, for the case of the optimal weights in a realistic survey, that approach is not feasible.

2.3 Calculation of the expectation of the galaxy density field

As a demonstration of how one can use the halo-galaxy double-delta expansion and take statistical averages we calculate the expectation of $\mathcal{F}_{\rm g}$. We use equation (8) to break $\langle\mathcal{F}_{\rm g}(\boldsymbol{r})\rangle$ into two parts:

$$\langle\mathcal{F}_{\rm g}(\boldsymbol{r})\rangle = \int_0^\infty {\rm d}L\Theta(\boldsymbol{r}|L)[\langle\mathcal{N}_{\rm g}(\boldsymbol{r},L)\rangle - \alpha\,\langle\mathcal{N}_{\rm s}(\boldsymbol{r},L)\rangle]\,, \tag{9}$$

where we introduced the weighted mean number density of galaxies per unit luminosity, at the spatial position $\boldsymbol{r}(\chi,\boldsymbol{\Omega})$:

$$\begin{aligned}\langle\mathcal{N}_{\rm g}(\boldsymbol{r},L)\rangle &\equiv \int {\rm d}^3x\int_0^\infty {\rm d}M\frac{w(\boldsymbol{r},L,\boldsymbol{x},M)}{\sqrt{A}}\langle n_{\rm g}(\boldsymbol{r},L,\boldsymbol{x},M)\rangle\\ &= \int {\rm d}^3x\int_0^\infty {\rm d}M\frac{w(\boldsymbol{r},L,\boldsymbol{x},M)}{\sqrt{A}}\left\langle\sum_{i=1}^{N_h}\delta^{\rm D}(\boldsymbol{x}-\boldsymbol{x}_i)\delta^{\rm D}(M-M_i)\sum_{j=1}^{N_{\rm g}(M_i)}\delta^{\rm D}(\boldsymbol{r}-\boldsymbol{r}_j-\boldsymbol{x}_i)\delta^{\rm D}(L-L_j)\right\rangle,\end{aligned} \tag{10}$$

with a similar expression for $\langle\mathcal{N}_{\rm s}(\boldsymbol{r},L)\rangle$. The function in equation (10) is related to the galaxy luminosity function.

To proceed further, we now need to understand what taking the ‘expectation value’ actually means. Following Smith (2012), this operation can be broken down into three steps. First, the fluctuations in the underlying dark matter field are sampled – we shall denote this averaging through a subscript s. Secondly, given the dark matter field, the haloes may be obtained as a sampling of the density field – we shall denote this operation through subscript h. Thirdly, given a set of dark matter haloes, and sufficient knowledge of the properties of the halo, galaxies may then be sampled into each halo – we shall denote this operation through subscript g. Hence, equation (10) can be rewritten:

$$\langle\mathcal{N}_{\rm g}(\boldsymbol{r},L)\rangle = \int {\rm d}^3x\int_0^\infty {\rm d}M\frac{w(\boldsymbol{r},L,\boldsymbol{x},M)}{\sqrt{A}}\left\langle\sum_{i=1}^{N_{\rm h}}\delta^{\rm D}(\boldsymbol{x}-\boldsymbol{x}_i)\delta^{\rm D}(M-M_i)\left\langle\sum_{j=1}^{N_{\rm g}(M_i)}\delta^{\rm D}(\boldsymbol{r}-\boldsymbol{r}_j-\boldsymbol{x}_i)\delta^{\rm D}(L-L_j)\right\rangle_{\rm g}\right\rangle_{\rm s,h}. \tag{11}$$

Let us now compute the average over the galaxy sampling for the ith dark matter halo:

$$\begin{aligned}\left\langle\sum_{j=1}^{N_{\rm g}(M_i)}\delta^{\rm D}(\boldsymbol{r}-\boldsymbol{r}_j-\boldsymbol{x}_i)\delta^{\rm D}(L-L_j)\right\rangle_{\rm g} &\equiv \sum_{N_{\rm g}=0}^{\infty}P(N_{\rm g}|\lambda(M_i))\int\prod_{k=1}^{N_{\rm g}}\left\{{\rm d}^3r_k{\rm d}L_k\right\}\,p(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_{N_{\rm g}},L_1,\ldots,L_{N_{\rm g}}|M_i,\boldsymbol{x}_i)\\ &\times\left[\delta^{\rm D}(\boldsymbol{r}-\boldsymbol{r}_1-\boldsymbol{x}_i)\delta^{\rm D}(L-L_1)+\ldots+\delta^{\rm D}(\boldsymbol{r}-\boldsymbol{r}_{N_{\rm g}}-\boldsymbol{x}_i)\delta^{\rm D}(L-L_{N_{\rm g}})\right],\end{aligned} \tag{12}$$

where in the above we have introduced the following quantities: $P(N_{\rm g}|\lambda(M_i))$ is the discrete probability that there are $N_{\rm g}$ galaxies in the ith dark matter halo and this we assume depends on some function of the dark matter halo mass M_i; $p(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_{N_{\rm g}},L_1,\ldots,L_{N_{\rm g}}|M_i,\boldsymbol{x}_i)$ is the joint probability density function for finding the $N_{\rm g}$ galaxies being located at positions $\{\boldsymbol{r}_1,\ldots,\boldsymbol{r}_{N_{\rm g}}\}$ relative to the halo centre $\boldsymbol{x}_i$, and with luminosities $\{L_1,\ldots,L_{N_{\rm g}}\}$, conditioned on M_i and $\boldsymbol{x}_i$. We have assumed that the properties and distribution of the galaxies in the ith halo are independent of all other external haloes. If the probability for finding a galaxy at a given position inside a halo is determined by the density profile of the matter in the halo, and if the probability that the galaxy has a luminosity L depends only on the halo mass, then this joint probability can be written in the following manner:

$$p(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_{N_{\rm g}},L_1,\ldots,L_{N_{\rm g}}|M_i,\boldsymbol{x}_i) = \prod_{k=1}^{N_{\rm g}}\{p(\boldsymbol{r}_k|M_i,\boldsymbol{x}_i)p(L_k|M_i)\} = \prod_{k=1}^{N_{\rm g}}\{U(\boldsymbol{r}_k|M_i,\boldsymbol{x}_i)\Phi(L_k|M_i)\}\,, \tag{13}$$

where in the above equation, we have used the density profile of galaxies in the halo, normalized by the total number of galaxies in that halo, U, to define

$$p(\boldsymbol{r}|M,\boldsymbol{x}) \equiv U(\boldsymbol{r}|M,\boldsymbol{x}) \equiv \rho_{\rm g}(\boldsymbol{r}|M,\boldsymbol{x})/N_{\rm g}(M)\,. \tag{14}$$

We have also used

$$p(L_k|M_i) \equiv \Phi(L_k|M_i)\,, \tag{15}$$

as the probability density that a galaxy hosted by a halo of mass M, has a luminosity L.[2] In writing equation (13) we have assumed that, for a given galaxy, its spatial location inside the dark matter halo is independent of its luminosity. As will be shown later, this assumption will not be crucial for the derivation of the optimal weights.

On integrating over the Dirac delta functions in equation (12) we find

$$\begin{aligned}\left\langle\sum_{j=1}^{N_{\rm g}(M_i)}\delta^{\rm D}(\boldsymbol{r}-\boldsymbol{r}_j-\boldsymbol{x}_i)\delta^{\rm D}(L-L_j)\right\rangle_{\rm g} &= \sum_{N_{\rm g}=0}^{\infty}P(N_{\rm g}|\lambda(M_i))N_{\rm g}U(\boldsymbol{r}-\boldsymbol{x}_i|M_i)\Phi(L|M_i)\\ &= N_{\rm g}^{(1)}(M_i)U(\boldsymbol{r}-\boldsymbol{x}_i|M_i)\Phi(L|M_i)\,,\end{aligned} \tag{16}$$

[2] Note that this is closely related to the conditional luminosity function introduced by Yang, Mo & van den Bosch (cf. 2003), which in our notation would be $\Phi_{\rm Yang\,et\,al}(L|M) = N_{\rm g}^{(1)}(M)\Phi(L|M)$.

where we have suppressed the dependence of U on the halo centre. The second equality follows from the definition of the first factorial moment of the galaxy distribution:

$$N_{\rm g}^{(1)}(M_i) \equiv \sum_{N_{\rm g}=0}^{\infty} P(N_{\rm g}|\lambda(M_i))N_{\rm g}\,. \tag{17}$$

Returning to our main calculation, on substituting the last two equations into equation (11), we now obtain

$$\begin{aligned}\langle \mathcal{N}_{\rm g}(\boldsymbol{r},L)\rangle &= \int {\rm d}^3x \int_0^{\infty} {\rm d}M \frac{w(\boldsymbol{r},L,\boldsymbol{x},M)}{\sqrt{A}} \left\langle \sum_{i=1}^{N_{\rm h}} \delta^{\rm D}(\boldsymbol{x}-\boldsymbol{x}_i)\delta^{\rm D}(M-M_i)N_{\rm g}^{(1)}(M_i)U(\boldsymbol{r}-\boldsymbol{x}_i|M_i)\Phi(L|M_i)\right\rangle_{\rm s,h}\\ &= \int {\rm d}^3x \int_0^{\infty} {\rm d}M \frac{w(\boldsymbol{r},L,\boldsymbol{x},M)}{\sqrt{A}} \int {\rm d}^3x_1 \dots {\rm d}^3x_{N_{\rm h}}{\rm d}M_1\dots{\rm d}M_{N_{\rm h}}\, p(\boldsymbol{x}_1,\dots \boldsymbol{x}_{N_{\rm h}},M_1,\dots,M_{N_{\rm h}})\\ &\quad \times \sum_{i=1}^{N_h} \delta^{\rm D}(\boldsymbol{x}-\boldsymbol{x}_i)\delta^{\rm D}(M-M_i)N_{\rm g}^{(1)}(M_i)U(\boldsymbol{r}-\boldsymbol{x}_i|M_i)\Phi(L|M_i)\,.\end{aligned} \tag{18}$$

In the above, we followed SM14 to introduce $p(\boldsymbol{x}_1,\dots,\boldsymbol{x}_{N_{\rm h}},M_1,\dots,M_{N_{\rm h}})$ as the joint probability density for the $N_{\rm h}$ dark matter halo centres being located at positions $\{\boldsymbol{x}_1,\dots,\boldsymbol{x}_{N_{\rm h}}\}$, and with masses $\{M_1,\dots,M_{N_{\rm h}}\}$. On integrating over the Dirac delta functions, the mean number density of galaxies becomes

$$\langle \mathcal{N}_{\rm g}(\boldsymbol{r},L)\rangle = \int {\rm d}^3x \int_0^{\infty} {\rm d}M \frac{w(\boldsymbol{r},L,\boldsymbol{x},M)}{\sqrt{A}} N_{\rm h}\, p(\boldsymbol{x},M) N_{\rm g}^{(1)}(M)U(\boldsymbol{r}-\boldsymbol{x}|M)\Phi(L|M)\,. \tag{19}$$

The joint distribution function for obtaining a halo of mass M at position $\boldsymbol{x}$ can be written as the product of two independent one-point probability density functions (Smith & Watts 2005):

$$p(\boldsymbol{x},M) = p(M)p(\boldsymbol{x}) = \frac{\bar{n}(M)}{\bar{n}_{\rm h}} \times \frac{1}{V_{\mu}} = \frac{\bar{n}(M)}{N_{\rm h}}\,, \tag{20}$$

where $\bar{n}(M)$ is the mean mass function of dark matter haloes, which tells us the number density of haloes of mass M, per unit mass, and $\bar{n}_{\rm h} = N_{\rm h}/V_{\mu}$ is the mean number density of haloes. On substituting this expression into equation (19), we find that the mean density of galaxies, per unit luminosity, at spatial location $\boldsymbol{r}$ may be written:

$$\langle \mathcal{N}_{\rm g}(\boldsymbol{r},L)\rangle = \frac{1}{\sqrt{A}}\phi_{\rm w}(\boldsymbol{r},L)\,, \tag{21}$$

where we have defined

$$\phi_{\rm w}(\boldsymbol{r},L) \equiv \int_0^{\infty} {\rm d}M\bar{n}(M)N_{\rm g}^{(1)}(M)\Phi(L|M) \int {\rm d}^3x\, w(\boldsymbol{r},L,\boldsymbol{x},M)U(\boldsymbol{r}-\boldsymbol{x}|M)\,. \tag{22}$$

If we were to set the weight function to unity, the above expression would be the galaxy luminosity function (Yang et al. 2003):

$$\phi(L) \equiv \int_0^{\infty} {\rm d}M\bar{n}(M)N_{\rm g}^{(1)}(M)\Phi(L|M)\,. \tag{23}$$

Turning to the second expectation value in equation (9), we note that the only difference between $\langle \mathcal{N}_{\rm g}(\boldsymbol{r},L)\rangle$ and $\langle \mathcal{N}_{\rm s}(\boldsymbol{r},L)\rangle$ is the artificially increased space density of clusters and the absence of any intrinsic clustering. Hence, we also have

$$\alpha\,\langle \mathcal{N}_{\rm s}(\boldsymbol{r},L)\rangle = \frac{1}{\sqrt{A}}\phi_{\rm w}(\boldsymbol{r},L)\,. \tag{24}$$

Returning to equation (9) and inserting equations (21) and (24) we arrive at the result:

$$\langle \mathcal{F}_{\rm g}(\boldsymbol{r})\rangle = 0\,. \tag{25}$$

Hence, the $\mathcal{F}_{\rm g}$–field, like the overdensity field of matter, is truly a mean-zero field.

Note that we have neglected to take into account the statistical properties of obtaining the $N_{\rm h}$ clusters in the survey volume. In what follows, we shall assume that the survey volumes are sufficiently large that this may be essentially treated as a deterministic quantity. However, it can be taken into account (e.g. see Sheth & Lemson 1999; Smith & Watts 2005; Smith 2009).

3 CLUSTERING ESTIMATORS

We now move on to the more interesting problem of using the halo-galaxy double-delta expansion to compute the clustering properties of the galaxy distribution. We begin first with the correlation function, and then through Fourier transforms look at the power spectrum. This task will be somewhat laborious; however, it will enable us to develop and establish a number of important concepts and results.

3.1 The two-point correlation function of galaxies

The two-point correlation function of the field $\mathcal{F}_{\rm g}$ can be computed using our double-delta expansion through:

$$\langle \mathcal{F}_{\rm g}(\boldsymbol{r}_1)\mathcal{F}_{\rm g}(\boldsymbol{r}_2)\rangle = \frac{1}{A}\int {\rm d}L_1{\rm d}L_2{\rm d}^3x_1{\rm d}^3x_2{\rm d}M_1{\rm d}M_2\Theta(\boldsymbol{r}_1|L_1)\Theta(\boldsymbol{r}_2|L_2)w(\boldsymbol{r}_1, L_1, \boldsymbol{x}_1, M_1)w(\boldsymbol{r}_2, L_2, \boldsymbol{x}_2, M_2)$$
$$\times [\, \langle n_{\rm g}(\boldsymbol{r}_1, L_1, \boldsymbol{x}_1, M_1)n_{\rm g}(\boldsymbol{r}_2, L_2, \boldsymbol{x}_2, M_2)\rangle - \alpha\langle n_{\rm g}(\boldsymbol{r}_1, L_1, \boldsymbol{x}_1, M_1)n_{\rm s}(\boldsymbol{r}_2, L_2, \boldsymbol{x}_2, M_2)\rangle$$
$$- \alpha\langle n_{\rm s}(\boldsymbol{r}_1, L_1, \boldsymbol{x}_1, M_1)n_{\rm g}(\boldsymbol{r}_2, L_2, \boldsymbol{x}_2, M_2)\rangle + \alpha^2\,\langle n_{\rm s}(\boldsymbol{r}_1, L_1, \boldsymbol{x}_1, M_1)n_{\rm s}(\boldsymbol{r}_2, L_2, \boldsymbol{x}_2, M_2)\rangle]\,. \quad (26)$$

The expectation terms in the square bracket on the right-hand side of this equation can be evaluated in a similar manner as was done for the case of the mean density. In Appendix A, we provide a detailed derivation of the terms $\langle n_{\rm g1}n_{\rm g2}\rangle$, $\langle n_{\rm s1}n_{\rm g2}\rangle$, $\langle n_{\rm g1}n_{\rm s2}\rangle$ and $\langle n_{\rm s1}n_{\rm s2}\rangle$. On substituting equations (A9), (A10), (A11) and (A12) into equation (26) and on integrating over the delta functions, we find that the correlation function may be written as the sum of three terms:

$$\langle \mathcal{F}_{\rm g}(\boldsymbol{r}_1)\mathcal{F}_{\rm g}(\boldsymbol{r}_2)\rangle = \prod_{i=1}^{2}\left\{\int {\rm d}^3x_i{\rm d}M_i\bar{n}(M_i)b(M_i)N_{\rm g}^{(1)}(M_i)\mathcal{W}_{(1)}^{U}(\boldsymbol{r}_i, \boldsymbol{x}_i, M_i)\right\}\xi(|\boldsymbol{x}_1-\boldsymbol{x}_2|)$$
$$+ (1+\alpha)\int {\rm d}^3x{\rm d}M\bar{n}(M)N_{\rm g}^{(2)}(M)\mathcal{W}_{(1)}^{U}(\boldsymbol{r}_1, \boldsymbol{x}, M)\mathcal{W}_{(1)}^{U}(\boldsymbol{r}_2, \boldsymbol{x}, M)$$
$$+ (1+\alpha)\int {\rm d}^3x{\rm d}M\bar{n}(M)N_{\rm g}^{(1)}(M)\mathcal{W}_{(2)}^{U}(\boldsymbol{r}_1, \boldsymbol{x}, M)\delta^{\rm D}(\boldsymbol{r}_1-\boldsymbol{r}_2)\,, \quad (27)$$

where $b(M)$ is the large-scale linear bias of dark matter haloes, $\xi(\boldsymbol{x})$ is the dark matter correlation function and $N_{\rm g}^{(2)}(M)$ is the second factorial moment of the galaxy numbers (for more details on these quantities see Appendix A). In the above expression we have also defined the quantity:

$$\mathcal{W}_{(l)}^{U}(\boldsymbol{r}, \boldsymbol{x}, M) \equiv U^{l}(\boldsymbol{r}-\boldsymbol{x}|M)\mathcal{W}_{(l)}(\boldsymbol{r}, \boldsymbol{x}, M)\,, \quad (28)$$

with

$$\mathcal{W}_{(l)}(\boldsymbol{r}, \boldsymbol{x}, M) \equiv \frac{1}{A^{l/2}}\int {\rm d}L\,\Phi(L|M)\Theta(\boldsymbol{r}|L)w^{l}(\boldsymbol{r}, L, \boldsymbol{x}, M)\,. \quad (29)$$

Based on the above analysis, we see that an obvious estimator for the $\mathcal{F}_{\rm g}$ correlation function is

$$\hat{\xi}_{\mathcal{F}_{\rm g}}(\boldsymbol{r}) \equiv \int {\rm d}^3r'\mathcal{F}_{\rm g}(\boldsymbol{r}')\mathcal{F}_{\rm g}(\boldsymbol{r}+\boldsymbol{r}') \quad ; \quad (\boldsymbol{r}\neq 0)\,. \quad (30)$$

The expectation of the estimator is:

$$\left\langle \hat{\xi}_{\mathcal{F}_{\rm g}}(\boldsymbol{r})\right\rangle = \int \prod_{i=1}^{2}\{{\rm d}^3x_i\,{\rm d}M_i\bar{n}(M_i)b(M_i)N_{\rm g}^{(1)}(M_i)\}\xi(|\boldsymbol{x}_1-\boldsymbol{x}_2|)\int {\rm d}^3r'\mathcal{W}_{(1)}^{U}(\boldsymbol{r}', \boldsymbol{x}_1, M_1)\mathcal{W}_{(1)}^{U}(\boldsymbol{r}+\boldsymbol{r}', \boldsymbol{x}_2, M_2)$$
$$+ (1+\alpha)\int {\rm d}^3x\,{\rm d}M\bar{n}(M)N_{\rm g}^{(2)}(M)\int {\rm d}^3r'\,\mathcal{W}_{(1)}^{U}(\boldsymbol{r}', \boldsymbol{x}, M)\,\mathcal{W}_{(1)}^{U}(\boldsymbol{r}+\boldsymbol{r}', \boldsymbol{x}, M) \quad ; \quad (\boldsymbol{r}\neq 0). \quad (31)$$

In general, $\hat{\xi}_{\mathcal{F}_{\rm g}}$ is a biased estimator for the matter correlation function ξ. Thus, in order to make a robust comparison between theory and observations, one must either compute the theory predictions as in equation (27) or generate Monte Carlo mock samples and use the same estimator to compare theory and observation.

3.2 An estimator for the matter correlation function in the large-scale limit

In the large-scale limit, the clustering of galaxies can be used to obtain an unbiased estimate of the matter correlation function. To see this note that, since the dark matter haloes in simulations are cuspy, on large scales the mass- or galaxy-number-normalized density profiles of galaxies behave approximately as Dirac delta functions. We shall therefore take

$$U^{\rm LS}(\boldsymbol{r}|M) \to \delta^{\rm D}(\boldsymbol{r})\,, \quad (32)$$

and on implementing this in equation (31), we find after integrating over the Dirac delta functions:

$$\left\langle \hat{\xi}_{\mathcal{F}_{\rm g}}(\boldsymbol{r})\right\rangle \approx \xi(\boldsymbol{r})\int \prod_{i=1}^{2}\{{\rm d}M_i\bar{n}(M_i)b(M_i)N_{\rm g}^{(1)}(M_i)\}\int {\rm d}^3r'\,\mathcal{W}_{(1)}(\boldsymbol{r}', \boldsymbol{r}', M_1)\,\mathcal{W}_{(1)}(\boldsymbol{r}+\boldsymbol{r}', \boldsymbol{r}+\boldsymbol{r}', M_2)\,;\ (\boldsymbol{r}\neq 0) \quad (33)$$

In this limit, there are a number of interesting things that happen: first, the weight function has now become independent of the halo positions, i.e. we no longer differentiate between galaxy and halo positions, taking them to be the same. Hence, we may now write:

$$w(\boldsymbol{r}, L, \boldsymbol{x}, M) \to w^{\rm LS}(\boldsymbol{r}, L, M)\ ;\ \mathcal{W}_{(l)}(\boldsymbol{r}, \boldsymbol{x}, M) \to \mathcal{W}_{(l)}^{\rm LS}(\boldsymbol{r}, M)\,. \quad (34)$$

Secondly, the dark matter correlation function has separated out and so we may easily invert equation (33) to obtain an unbiased estimate for the dark matter clustering. The estimator is:

$$\hat{\xi}(\boldsymbol{r}) \approx \frac{\hat{\xi}_{\mathcal{F}_{\rm g}}(\boldsymbol{r})}{\Sigma_0(\boldsymbol{r})} \quad ; \quad \Sigma_0(\boldsymbol{r}) \equiv \int {\rm d}^3 r' \, \mathcal{G}^{(1)}_{(1,1)}(\boldsymbol{r}') \, \mathcal{G}^{(1)}_{(1,1)}(\boldsymbol{r}+\boldsymbol{r}') \quad ; \quad (\boldsymbol{r} \neq 0) \tag{35}$$

where we have defined the new set of weighted window functions:

$$\mathcal{G}^{(n)}_{(l,m)}(\boldsymbol{r}) \equiv \int {\rm d}M \bar{n}(M) b^m(M) N^{(n)}_{\rm g}(M) \left[\mathcal{W}^{\rm LS}_{(l)}(\boldsymbol{r}, M)\right]^n . \tag{36}$$

The function $\Sigma_0(\boldsymbol{r})$ represents the correlation function of the averaged survey window functions. Our correlation function estimator equation (35), is therefore a generalization of that of Landy & Szalay (1993).

3.3 The galaxy power spectrum

We now turn to the Fourier space dual of the correlation function – the galaxy power spectrum. To obtain this let us begin by defining our 3D Fourier transform convention for a function B and its inverse as:

$$\tilde{B}(\boldsymbol{k}) \equiv \int {\rm d}^3 r \, B(\boldsymbol{r}) {\rm e}^{i\,\boldsymbol{k}\cdot\boldsymbol{r}} \qquad \Leftrightarrow \qquad B(\boldsymbol{r}) = \int \frac{{\rm d}^3 k}{(2\pi)^3} \tilde{B}(\boldsymbol{k}) {\rm e}^{-i\,\boldsymbol{k}\cdot\boldsymbol{r}} .$$

We distinguish real- and Fourier-space quantities that share the same symbol through use of the tilde notation. We also define the power spectrum $P_B(k)$ of any infinite statistically homogeneous random field $\tilde{B}(\boldsymbol{k})$ to be:

$$\langle \tilde{B}(\boldsymbol{k}) \tilde{B}(\boldsymbol{k}') \rangle \equiv (2\pi)^3 \delta^{\rm D}(\boldsymbol{k}+\boldsymbol{k}') P_B(\boldsymbol{k}) .$$

Note that if the field B were statistically isotropic, the power spectrum would simply be a function of the scalar k. In addition, the power spectrum and two-point correlation function of the field B form a Fourier pair:

$$\xi_B(|\boldsymbol{x}-\boldsymbol{x}'|) = \int \frac{{\rm d}^3 q}{(2\pi)^3} \int \frac{{\rm d}^3 q'}{(2\pi)^3} P_B(\boldsymbol{q}) (2\pi)^3 \delta^{\rm D}(\boldsymbol{q}+\boldsymbol{q}') {\rm e}^{-i\boldsymbol{q}\cdot\boldsymbol{x}} {\rm e}^{-i\boldsymbol{q}'\cdot\boldsymbol{x}'} .$$

With these definitions in hand we may now transform equation (27), and on considering the case where $\boldsymbol{k}_2 = -\boldsymbol{k}$, we find that the expectation of the square of the amplitude of the Fourier modes of $\tilde{\mathcal{F}}_{\rm g}$ is given by

$$\begin{aligned}\left\langle |\tilde{\mathcal{F}}_{\rm g}(\boldsymbol{k})|^2 \right\rangle = & \int \frac{{\rm d}^3 q}{(2\pi)^3} P(\boldsymbol{q}) \prod_{i=1}^{2} \left\{ \int {\rm d}M_i \bar{n}(M_i) b(M_i) N^{(1)}_{\rm g}(M_i) \right\} \tilde{\mathcal{W}}^U_{(1)}(\boldsymbol{k}, -\boldsymbol{q}, M_1) \tilde{\mathcal{W}}^U_{(1)}(-\boldsymbol{k}, \boldsymbol{q}, M_2) \\ & + (1+\alpha) \int \frac{{\rm d}^3 q}{(2\pi)^3} {\rm d}M \, \bar{n}(M) N^{(2)}_{\rm g}(M) \tilde{\mathcal{W}}^U_{(1)}(\boldsymbol{k}, -\boldsymbol{q}, M) \tilde{\mathcal{W}}^U_{(1)}(-\boldsymbol{k}, \boldsymbol{q}, M) \\ & + (1+\alpha) \int {\rm d}^3 r {\rm d}^3 x \; \bar{n}(M) N^{(1)}_{\rm g}(M) \tilde{\mathcal{W}}^U_{(2)}(\boldsymbol{r}, \boldsymbol{x}, M) ,\end{aligned} \tag{37}$$

where $P(\boldsymbol{q})$ is the matter power spectrum and we have set the Fourier transform of the effective survey window function to be

$$\tilde{\mathcal{W}}^U_{(l)}(\boldsymbol{k}, \boldsymbol{q}, M) \equiv \int {\rm d}^3 r {\rm d}^3 x \, \mathcal{W}^U_{(l)}(\boldsymbol{r}, \boldsymbol{x}, M) {\rm e}^{i\boldsymbol{k}\cdot\boldsymbol{r}} {\rm e}^{i\boldsymbol{q}\cdot\boldsymbol{x}} . \tag{38}$$

As in the case of the correlation function of the field $\mathcal{F}_{\rm g}$, its power spectrum does not provide a direct estimate of the matter power spectrum.

3.4 The power spectrum in the large-scale limit

Let us now consider the power spectrum of $\tilde{\mathcal{F}}_{\rm g}$ in the large-scale limit. As discussed in Section 3.2 we expect the density profiles to behave like Dirac delta functions in real space, in Fourier space the density profiles on large scales simply obey: $\tilde{U}(\boldsymbol{k}|M) \to k \to 01$. Under this condition equation (37) simplifies to

$$\left\langle |\tilde{\mathcal{F}}_{\rm g}(\boldsymbol{k})|^2 \right\rangle \approx \int \frac{{\rm d}^3 q}{(2\pi)^3} P(\boldsymbol{q}) \left| \tilde{\mathcal{G}}^{(1)}_{(1,1)}(\boldsymbol{k}-\boldsymbol{q}) \right|^2 + P_{\rm shot}, \tag{39}$$

where the second term on the right-hand side is a k-independent effective shot-noise term,

$$P_{\rm shot} \equiv (1+\alpha) \left[\tilde{\mathcal{G}}^{(2)}_{(1,0)}(\boldsymbol{0}) + \tilde{\mathcal{G}}^{(1)}_{(2,0)}(\boldsymbol{0}) \right] . \tag{40}$$

In the limit that the survey volume is large, the window functions $\tilde{\mathcal{G}}^{(n)}_{(l,m)}(\boldsymbol{k})$ will be very narrowly peaked around $\boldsymbol{k} = \boldsymbol{0}$. Provided the matter power spectrum is a smoothly varying function of scale, the window functions $\tilde{\mathcal{G}}^{(n)}_{(l,m)}(\boldsymbol{k})$ will behave in a way that is similar to the Dirac delta function. Hence, equation (39) becomes

$$\left\langle |\tilde{\mathcal{F}}_{\rm g}(\boldsymbol{k})|^2 \right\rangle \approx P(\boldsymbol{k}) \int \frac{{\rm d}^3 q}{(2\pi)^3} \left| \tilde{\mathcal{G}}^{(1)}_{(1,1)}(\boldsymbol{k}-\boldsymbol{q}) \right|^2 + P_{\rm shot}. \tag{41}$$

Let us focus on the integral factor on the right-hand side of the above expression. If we now perform the transformation of variables $\boldsymbol{q} \to \boldsymbol{k} - \boldsymbol{q}$ and use Parseval's theorem, we find

$$\int \frac{\mathrm{d}^3 q}{(2\pi)^3} \left| \tilde{\mathcal{G}}^{(1)}_{(1,1)}(\boldsymbol{k}-\boldsymbol{q}) \right|^2 = \int \frac{\mathrm{d}^3 q}{(2\pi)^3} \left| \tilde{\mathcal{G}}^{(1)}_{(1,1)}(\boldsymbol{q}) \right|^2 = \int \mathrm{d}^3 r \left| \mathcal{G}^{(1)}_{(1,1)}(\boldsymbol{r}) \right|^2 .$$

Upon back substitution of equation (36) into the above expression, we obtain:

$$\int \mathrm{d}^3 r \left| \mathcal{G}^{(1)}_{(1,1)}(\boldsymbol{r}) \right|^2 = \frac{1}{A} \int \mathrm{d}^3 r \left[\int \mathrm{d}M \bar{n}(M) N^{(1)}_{\mathrm{g}}(M) b(M) \int \mathrm{d}L\, \Phi(L|M) \Theta(\boldsymbol{r}|L) w(\boldsymbol{r}, L, M) \right]^2 .$$

Note that we have not yet specified the parameter A, which we now take to be:

$$A \equiv \int \mathrm{d}^3 r \left[\int \mathrm{d}M \bar{n}(M) N^{(1)}_{\mathrm{g}}(M) b(M) \int \mathrm{d}L\, \Phi(L|M) \Theta(\boldsymbol{r}|L) w(\boldsymbol{r}, L, M) \right]^2 . \tag{42}$$

Note that we will now drop the superscript LS notation for w and $\mathcal{W}_{(l)}$, since for the remainder of this study we shall be working only in the large-scale limit. Thus, our estimator for the matter power spectrum can be written simply

$$\hat{P}(\boldsymbol{k}) = |\tilde{\mathcal{F}}_{\mathrm{g}}(\boldsymbol{k})|^2 - P_{\mathrm{shot}} . \tag{43}$$

If our modelling of the galaxy distribution is correct, then the above estimator constitutes the only unbiased estimator of the matter power spectrum. Before proceeding further, note that in the above expression, we have obtained the power spectrum per mode. In fact, we are more interested in its band-power estimate. Hence, our final estimator in the large-scale and large-volume limit is

$$\overline{P}(k_i) = \frac{1}{V_i} \int_{V_i} \mathrm{d}^3 k\, \hat{P}(\boldsymbol{k}) = \frac{1}{V_i} \int_{V_i} \mathrm{d}^3 k |\tilde{\mathcal{F}}_{\mathrm{g}}(\boldsymbol{k})|^2 - P_{\mathrm{shot}} , \tag{44}$$

where in the above we have summed over all modes in a k-space shell of thickness Δk and volume

$$V_i \equiv \int_{V_i} \mathrm{d}^3 k = 4\pi \int_{k_i - \Delta k/2}^{k_i + \Delta k/2} k^2 \mathrm{d}k = 4\pi k_i \Delta k \left[1 + \frac{1}{12} \left(\frac{\Delta k}{k_i} \right)^2 \right] . \tag{45}$$

4 STATISTICAL FLUCTUATIONS IN THE GALAXY POWER SPECTRUM

In order to obtain the optimal estimator we need to know how the signal to noise (hereafter $\mathcal{S}/\mathcal{N}$) varies when we vary the shape of our weight function w. Thus, we need to understand the noise properties of our power spectrum estimator, i.e. compute its covariance matrix. The covariance matrix of two band-power estimates is given by

$$\mathrm{Cov}\left[\overline{P}(k_i), \overline{P}(k_j)\right] \equiv \left\langle \overline{P}(k_i) \overline{P}(k_j) \right\rangle - \left\langle \overline{P}(k_i) \right\rangle \left\langle \overline{P}(k_j) \right\rangle = \frac{1}{V_i} \int_{V_i} \mathrm{d}^3 k_1 \frac{1}{V_j} \int_{V_j} \mathrm{d}^3 k_2\, \mathrm{Cov}\left[\hat{P}(\boldsymbol{k}_1), \hat{P}(\boldsymbol{k}_2) \right] ,$$

where the last factor on the right-hand side of the above expression is the covariance of the power in two separate Fourier modes. For the case of large survey volumes and in the large-scale limit, the matter power spectrum is given by equation (44). Hence,

$$\mathrm{Cov}\left[\hat{P}(\boldsymbol{k}_1), \hat{P}(\boldsymbol{k}_2) \right] \approx \mathrm{Cov}\left[|\tilde{\mathcal{F}}_{\mathrm{g}}(\boldsymbol{k}_1)|^2, |\tilde{\mathcal{F}}_{\mathrm{g}}(\boldsymbol{k}_2)|^2 \right] = \left\langle |\tilde{\mathcal{F}}_{\mathrm{g}}(\boldsymbol{k}_1)|^2 |\tilde{\mathcal{F}}_{\mathrm{g}}(\boldsymbol{k}_2)|^2 \right\rangle - \left\langle |\tilde{\mathcal{F}}_{\mathrm{g}}(\boldsymbol{k}_1)|^2 \right\rangle \left\langle |\tilde{\mathcal{F}}_{\mathrm{g}}(\boldsymbol{k}_2)|^2 \right\rangle . \tag{46}$$

The approximation in the equation above follows from the discussion in Appendix B1.

In Appendix C, we derive a general expression for the covariance matrix of $|\tilde{\mathcal{F}}_{\mathrm{g}}(\boldsymbol{k}_1)|^2$ and $|\tilde{\mathcal{F}}_{\mathrm{g}}(\boldsymbol{k}_2)|^2$, with all n-point connected spectra and shot-noise terms included – this is obtained by combining equations (C1) and (C2) with equation (C13). Under the assumption of a Gaussian matter density field, our general expression simplifies to equation (C18). Furthermore, we also show in Appendix C3 that in the large-scale limit equation (C18) can be written as

$$\begin{aligned} \mathrm{Cov}\left[|\tilde{\mathcal{F}}_{\mathrm{g}}(\boldsymbol{k}_1)|^2, |\tilde{\mathcal{F}}_{\mathrm{g}}(\boldsymbol{k}_2)|^2 \right] = & \left| \int \frac{\mathrm{d}^3 q}{(2\pi)^3} P(\boldsymbol{q}) \tilde{\mathcal{G}}^{(1)}_{(1,1)}(\boldsymbol{k}_1 + \boldsymbol{q}) \tilde{\mathcal{G}}^{(1)}_{(1,1)}(\boldsymbol{k}_2 - \boldsymbol{q}) + (1+\alpha) \left[\tilde{\mathcal{G}}^{(1)}_{(2,0)}(\boldsymbol{k}_1 + \boldsymbol{k}_2) + \tilde{\mathcal{G}}^{(2)}_{(1,0)}(\boldsymbol{k}_1 + \boldsymbol{k}_2) \right] \right|^2 \\ & + \left| \int \frac{\mathrm{d}^3 q}{(2\pi)^3} P(\boldsymbol{q}) \tilde{\mathcal{G}}^{(1)}_{(1,1)}(\boldsymbol{k}_1 + \boldsymbol{q}) \tilde{\mathcal{G}}^{(1)}_{(1,1)}(-\boldsymbol{k}_2 - \boldsymbol{q}) + (1+\alpha) \left[\tilde{\mathcal{G}}^{(1)}_{(2,0)}(\boldsymbol{k}_1 - \boldsymbol{k}_2) + \tilde{\mathcal{G}}^{(2)}_{(1,0)}(\boldsymbol{k}_1 - \boldsymbol{k}_2) \right] \right|^2 . \end{aligned} \tag{47}$$

In the limit where the survey volume is large, the functions $\tilde{\mathcal{G}}^{(n)}_{(l,m)}$ are very narrowly peaked around $k = 0$. Furthermore, if the power spectrum does not rapidly vary over the scale of the effective window function, then we may treat it as a constant in equation (47). Thus,

$$\begin{aligned} \mathrm{Cov}\left[|\tilde{\mathcal{F}}_{\mathrm{g}}(\boldsymbol{k}_1)|^2, |\tilde{\mathcal{F}}_{\mathrm{g}}(\boldsymbol{k}_2)|^2 \right] \approx & \left| P(\boldsymbol{k}_1) \tilde{\mathcal{Q}}^{(1,1)}_{(1,1|1,1)}(\boldsymbol{k}_1 + \boldsymbol{k}_2) + (1+\alpha) \left[\tilde{\mathcal{Q}}^{(2)}_{(1|0)}(\boldsymbol{k}_1 + \boldsymbol{k}_2) + \tilde{\mathcal{Q}}^{(1)}_{(2|0)}(\boldsymbol{k}_1 + \boldsymbol{k}_2) \right] \right|^2 \\ & + \left| P(\boldsymbol{k}_1) \tilde{\mathcal{Q}}^{(1,1)}_{(1,1|1,1)}(\boldsymbol{k}_1 - \boldsymbol{k}_2) + (1+\alpha) \left[\tilde{\mathcal{Q}}^{(2)}_{(1|0)}(\boldsymbol{k}_1 - \boldsymbol{k}_2) + \tilde{\mathcal{Q}}^{(1)}_{(2|0)}(\boldsymbol{k}_1 - \boldsymbol{k}_2) \right] \right|^2 , \end{aligned} \tag{48}$$

where we have introduced the functions:

$$\mathcal{Q}^{(n_1,n_2)}_{(l_1,l_2|m_1,m_2)}(\boldsymbol{r}) \equiv \mathcal{G}^{(n_1)}_{(l_1,m_1)}(\boldsymbol{r}) \mathcal{G}^{(n_2)}_{(l_2,m_2)}(\boldsymbol{r}) , \tag{49}$$

and made use of the convolution theorem to write their Fourier transforms:

$$\mathcal{Q}^{(n_1,n_2)}_{(l_1,l_2|m_1,m_2)}(\boldsymbol{k}) = \int \frac{\mathrm{d}^3 q}{(2\pi)^3} \mathcal{G}^{(n_1)}_{(l_1,m_1)}(\boldsymbol{q}) \mathcal{G}^{(n_2)}_{(l_2,m_2)}(\boldsymbol{k}-\boldsymbol{q})\,.$$

Note that we also used the trivial identity $\mathcal{Q}^{(n)}_{(l|m)} = \mathcal{G}^{(n)}_{(l,m)}$.

Returning to equation (46), we find that after substitution of equation (48) into equation (46), the bin-averaged estimates of the power spectrum can be written:

$$\mathrm{Cov}\left[\overline{P}(k_i), \overline{P}(k_j)\right] = 2\int_{V_i} \frac{\mathrm{d}^3 k_1}{V_i} \int_{V_j} \frac{\mathrm{d}^3 k_2}{V_j} \left| P(\boldsymbol{k}_1) \tilde{\mathcal{Q}}^{(1,1)}_{(1,1|1,1)}(\boldsymbol{k}_1 + \boldsymbol{k}_2) + (1+\alpha)\left[\tilde{\mathcal{Q}}^{(2)}_{(1|0)}(\boldsymbol{k}_1+\boldsymbol{k}_2) + \tilde{\mathcal{Q}}^{(1)}_{(2|0)}(\boldsymbol{k}_1+\boldsymbol{k}_2)\right]\right|^2; \tag{50}$$

equation (50) follows from the integrals over $\boldsymbol{k}_2$ in equation (48) being invariant under the transformation $\boldsymbol{k}_2 \to -\boldsymbol{k}_2$. Furthermore, if the k-space shells are narrow compared to the scale over which the power spectrum varies, then the shell-averaged power spectrum can be pulled out of the integrals. In Appendix D, we detail the computation of equation (50) and show that the covariance can be reexpressed as:

$$\mathrm{Cov}\left[\overline{P}(k_i), \overline{P}(k_j)\right] = \frac{2(2\pi)^3}{V_i} \overline{P}^2(k_i)\delta^K_{i,j} \int \mathrm{d}^3 r \left\{\left[\mathcal{G}^{(1)}_{(1,1)}(\boldsymbol{r})\right]^2 + \frac{(1+\alpha)}{\overline{P}(k_i)}\left[\mathcal{G}^{(2)}_{(1,0)}(\boldsymbol{r}) + \mathcal{G}^{(1)}_{(2,0)}(\boldsymbol{r})\right]\right\}^2, \tag{51}$$

with the functions $\mathcal{G}^{(n)}_{(l,m)}(\boldsymbol{r})$ defined by equation (36).

5 OPTIMAL ESTIMATOR

Our aim is to find the optimal weighting scheme that will maximize the $\mathcal{S}/\mathcal{N}$ ratio on a given band-power estimate of the galaxy power spectrum.

5.1 The optimal weight equation

To begin, note that maximizing the $\mathcal{S}/\mathcal{N}$ ratio is equivalent to minimizing its inverse, the noise-to-signal ratio $\mathcal{N}/\mathcal{S}$. The square of the latter can be expressed as:

$$F[w(\boldsymbol{r}, L, M)] \equiv \frac{\sigma^2_{\overline{P}}(k_i)}{\overline{P}^2(k_i)} = \frac{2(2\pi)^3}{V_i} \int \mathrm{d}^3 r \left\{\left[\mathcal{G}^{(1)}_{(1,1)}(\boldsymbol{r})\right]^2 + \frac{(1+\alpha)}{\overline{P}(k_i)}\left[\mathcal{G}^{(2)}_{(1,0)}(\boldsymbol{r}) + \mathcal{G}^{(1)}_{(2,0)}(\boldsymbol{r})\right]\right\}^2. \tag{52}$$

In the above expression, we have written the squared noise-to-signal $F[w]$ as a functional of the weights $w(\boldsymbol{r}, L, M)$. The standard way for finding the optimal weights is to perform the variation of the functional F with respect to the weights $w(\boldsymbol{r}, L, M)$. Operationally, the functional variation of $F[w]$ is carried out by comparing $F[w]$ with the functional obtained for weight functions that possess a small path variation $w(\boldsymbol{r}, L, M) \to w(\boldsymbol{r}, L, M) + \delta w(\boldsymbol{r}, L, M)$. This variation can be defined:

$$\delta F[w] \equiv F[w(\boldsymbol{r}, L, M) + \delta w(\boldsymbol{r}, L, M)] - F[w(\boldsymbol{r}, L, M)] = \int \mathrm{d}^3 r \mathrm{d}L \mathrm{d}M \left\{\frac{\delta F}{\delta w(\boldsymbol{r}, L, M)}\right\} \delta w(\boldsymbol{r}, L, M)\,. \tag{53}$$

Extremization means that the functional derivative is stationary for small variations around the optimal weights:

$$\frac{\delta F}{\delta w(\boldsymbol{r}, L, M)} = 0\,. \tag{54}$$

Recall that the definition of the weights in equation (29) includes the normalization constant A specified by equation (42). Since the normalization constant A is itself a function of the weights, it follows that $F[w]$ is in fact a ratio of two weight-dependent functionals:

$$F[w] \equiv \frac{\mathcal{N}[w]}{\mathcal{D}[w]}\,, \tag{55}$$

with the definitions:

$$\mathcal{N}[w] \equiv \int \mathrm{d}^3 r \left\{\left[\overline{\mathcal{G}}^{(1)}_{(1,1)}(\boldsymbol{r})\right]^2 + c\left[\overline{\mathcal{G}}^{(2)}_{(1,0)}(\boldsymbol{r}) + \overline{\mathcal{G}}^{(1)}_{(2,0)}(\boldsymbol{r})\right]\right\}^2; \tag{56}$$

$$\mathcal{D}[w] \equiv A^2[w] = \left[\int \mathrm{d}^3 r \left[\overline{\mathcal{G}}^{(1)}_{(1,1)}(\boldsymbol{r})\right]^2\right]^2. \tag{57}$$

In the above, we introduced the scaled effective window functions,

$$\overline{\mathcal{G}}^{(n)}_{(l,m)}(\boldsymbol{r}) = A^{nl/2}\mathcal{G}^{(n)}_{(l,m)}(\boldsymbol{r})\,, \tag{58}$$

as well as the constant $c \equiv (1+\alpha)/\overline{P}(k_i)$, which helps keep the equations as compact as possible. We also dropped the overall constant $2(2\pi)^3/V_i$ from the functional $\mathcal{N}[w]$, since it plays no role in the minimization process.

Minimizing $F[w]$ is equivalent to solving the functional problem:

$$\frac{1}{\mathcal{D}[w]}\left(\delta\mathcal{N}[w]-\frac{\mathcal{N}[w]}{\mathcal{D}[w]}\delta\mathcal{D}[w]\right)=0 \iff \delta\mathcal{N}[w]-F[w]\delta\mathcal{D}[w]=0. \tag{59}$$

Therefore, to find the optimal weights satisfying equation (59), we first need to compute the variations of $\mathcal{N}$ and $\mathcal{D}$ with a perturbation δw. This calculation is outlined in Appendix E. Putting together equations (59), (E6), (E7), we arrive at the following general equation for the optimal weights:

$$\left\{[\overline{\mathcal{G}}^{(1)}_{(1,1)}(\boldsymbol{r})]^2+c\left[\overline{\mathcal{G}}^{(2)}_{(1,0)}(\boldsymbol{r})+\overline{\mathcal{G}}^{(1)}_{(2,0)}(\boldsymbol{r})\right]\right\}\left\{\overline{\mathcal{G}}^{(1)}_{(1,1)}(\boldsymbol{r})b(M)+c\left[w(\boldsymbol{r},L,M)+\overline{\mathcal{W}}_1(\boldsymbol{r},M)\beta(M)N^{(1)}_{\rm g}(M)\right]\right\}=\overline{\mathcal{G}}^{(1)}_{(1,1)}(\boldsymbol{r})\,b(M)\,. \tag{60}$$

In the above, $\overline{\mathcal{W}}_1$ was introduced by equation (E4), and the function $\beta(M)$ specifies the relation between the first and second factorial moments of galaxies in a halo of mass M, as discussed in Cooray & Sheth (2002):

$$N^{(2)}_{\rm g}(M)=\beta(M)\left[N^{(1)}_{\rm g}(M)\right]^2\,. \tag{61}$$

Note that for a Poisson distribution, $\beta=1$, although we do not make this assumption here.

On inspection of equation (60) we notice that, with the exception of the weights $w(\boldsymbol{r}, L, M)$, none of the terms carries any explicit dependence on the luminosity of the galaxies. We therefore conclude that *the optimal weights are independent of luminosity*. Hence, without any loss of generality, we may now redefine the weights to be:

$$w(\boldsymbol{r},L,M)\Rightarrow w(\boldsymbol{r},M)\,. \tag{62}$$

One immediate consequence of this is that the functions $\overline{\mathcal{W}}_{(l)}(\boldsymbol{r}, M)$ can now be written in the much simplified form:

$$\overline{\mathcal{W}}_{(l)}(\boldsymbol{r},M)=w^l(\boldsymbol{r},M)\int {\rm d}L\,\Phi(L|M)\Theta(\boldsymbol{r}|L)=w^l(\boldsymbol{r},M)\mathcal{S}(\boldsymbol{r},M)\,.$$

In the above we have introduced the function:

$$\mathcal{S}(\boldsymbol{r},M)\equiv\Theta(\boldsymbol{\Omega})\int_0^\infty {\rm d}L\,\Theta(\chi|L)\Phi(L|M)=\Theta(\boldsymbol{\Omega})\int_{L_{\rm min}(\chi)}^\infty {\rm d}L\,\Phi(L|M)\,, \tag{63}$$

with the second equality following from equation (6). $\mathcal{S}(\boldsymbol{r}, M)$ is the number of galaxies in a halo of mass M observable at comoving distance $\boldsymbol{r}$ relative to the total number of galaxies in that halo. The range of S is the interval [0, 1], and it has the following limiting behaviour: for $M \ge M_{\rm min}$ we have $\lim_{\chi\to 0}\mathcal{S}(\boldsymbol{r}, M) = 1$ and $\lim_{\chi\to\infty}\mathcal{S}(\boldsymbol{r}, M) = 0$; and for $M < M_{\rm min}$ we have $\mathcal{S}(\boldsymbol{r}, M) = 0$, where $M_{\rm min}$ is the minimum halo mass required for a dark matter halo to be able to host a galaxy. Note that for a volume-limited survey $\mathcal{S}$ = constant.

A further consequence of equation (62) is that the effective survey window functions given by equation (58) reduce to:

$$\overline{\mathcal{G}}^{(n)}_{(l,m)}(\boldsymbol{r})\equiv\int {\rm d}M\bar{n}(M)b^m(M)N^{(n)}_{\rm g}(M)\left[w^l(\boldsymbol{r},M)\mathcal{S}(\boldsymbol{r},M)\right]^n\,. \tag{64}$$

Implementing these considerations in equation (60), we arrive at the equation governing the optimal weights:

$$\left\{\left[\overline{\mathcal{G}}^{(1)}_{(1,1)}(\boldsymbol{r})\right]^2+c\left[\overline{\mathcal{G}}^{(2)}_{(1,0)}(\boldsymbol{r})+\overline{\mathcal{G}}^{(1)}_{(2,0)}(\boldsymbol{r})\right]\right\}\left\{\overline{\mathcal{G}}^{(1)}_{(1,1)}(\boldsymbol{r})+c\frac{w(\boldsymbol{r},M)}{b(M)}\left[1+\beta(M)N^{(1)}_{\rm g}(M)\mathcal{S}(\boldsymbol{r},M)\right]\right\}=\overline{\mathcal{G}}^{(1)}_{(1,1)}(\boldsymbol{r})\,. \tag{65}$$

5.2 The optimal weights

We now seek a general solution for the weight equation (65). To begin, we notice that the only part of the weight equation that carries any mass dependence is the second bracket on the left-hand side of equation (65). If we set the radial vector to a constant $\boldsymbol{r} = \boldsymbol{r}_0$, then the optimal weights at fixed position inside the angular mask must have the mass dependence:

$$w(\boldsymbol{r}_0,M)\propto\frac{b(M)}{1+\beta(M)N^{(1)}_{\rm g}(M)\mathcal{S}(\boldsymbol{r}_0,M)}\,. \tag{66}$$

The weights are therefore proportional to the bias of the dark matter halo in which the galaxy is hosted and inversely proportional to the factor $[1+\beta(M)N^{(1)}_{\rm g}(M)\mathcal{S}(\boldsymbol{r}_0, M)]$. Since this last term depends on the galaxy selection function $\mathcal{S}$, the weight function is not separable in position and mass, as was found by SM14 for the case of optimal weighting of a sample of galaxy clusters. Nevertheless, without any loss of generality, we can factor out this part of the weight function from the general weight solution:

$$w(\boldsymbol{r},M)=\tilde{w}(\boldsymbol{r})\left[\frac{b(M)}{1+\beta(M)N^{(1)}_{\rm g}(M)(\boldsymbol{r},M)}\right]\,, \tag{67}$$

where $\tilde{w}(\boldsymbol{r})$ is a function of position only that needs to be determined. It is clear from the above equation that the term on the right-hand side encompasses the whole mass dependence of the optimal weights. If we now reinsert this expression into equation (65) we see that the weight equation reduces to:

$$\left\{\left[\overline{\mathcal{G}}^{(1)}_{(1,1)}(\boldsymbol{r})\right]^2+c\left[\overline{\mathcal{G}}^{(2)}_{(1,0)}(\boldsymbol{r})+\overline{\mathcal{G}}^{(1)}_{(2,0)}(\boldsymbol{r})\right]\right\}\left\{\overline{\mathcal{G}}^{(1)}_{(1,1)}(\boldsymbol{r})+c\tilde{w}(\boldsymbol{r})\right\}=\overline{\mathcal{G}}^{(1)}_{(1,1)}(\boldsymbol{r})\,. \tag{68}$$

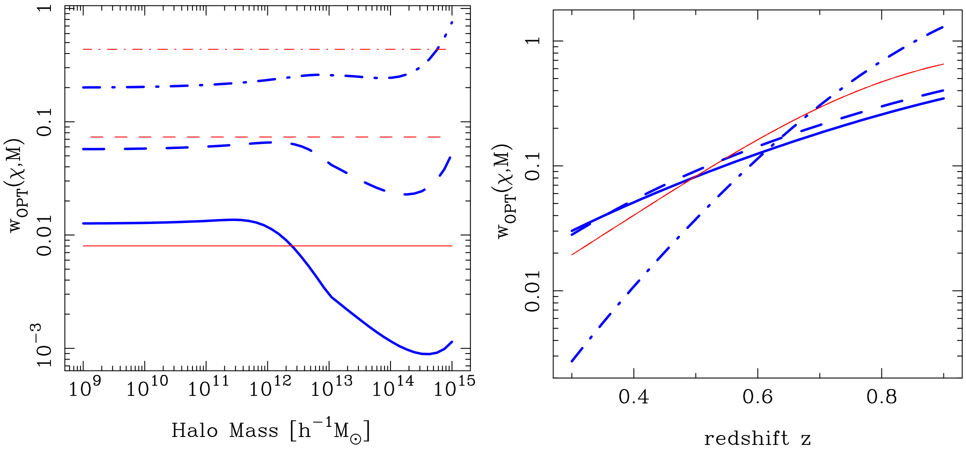

Figure 1. Left-hand panel: evolution of the optimal weights in a fiducial flux-limited survey as a function of halo mass. The thick blue and thin red lines represent the optimal weights and the FKP weights, respectively. The solid, dashed and dot–dashed line styles denote the results for increasing χ, respectively. Right-hand panel: evolution of the optimal weights as a function of redshift for several halo masses. Thick blue and thin red lines denote optimal and FKP weights. The solid, dashed and dot–dashed lines show the results for galaxy-, group- and cluster-scale halo masses, respectively. We have taken the flux limit to be $m_{\rm lim} = 22$.

In order to proceed further, we need to recompute the effective survey window functions $\overline{\mathcal{G}}$ functions from equation (64) with the new weight function equation (67). It is straightforward to show that

$$\overline{\mathcal{G}}^{(1)}_{(1,1)}(\boldsymbol{r}) = \tilde{w}(\boldsymbol{r})\bar{n}_{\rm eff}(\boldsymbol{r})\,, \text{ where } \ n_{\rm eff}(\boldsymbol{r}) \equiv \int {\rm d}M\bar{n}(M)b^2(M)\left[\frac{N^{(1)}_{\rm g}(M)S(\boldsymbol{r},M)}{1+\beta(M)N^{(1)}_{\rm g}(M)S(\boldsymbol{r},M)}\right]; \tag{69}$$

$$\overline{\mathcal{G}}^{(2)}_{(1,0)}(\boldsymbol{r}) = \tilde{w}^2(\boldsymbol{r})\int {\rm d}M\bar{n}(M)b^2(M)\beta(M)\left[\frac{N^{(1)}_{\rm g}(M)S(\boldsymbol{r},M)}{1+\beta(M)N^{(1)}_{\rm g}(M)S(\boldsymbol{r},M)}\right]^2; \tag{70}$$

$$\overline{\mathcal{G}}^{(1)}_{(2,0)}(\boldsymbol{r}) = \tilde{w}^2(\boldsymbol{r})\int {\rm d}M\bar{n}(M)b^2(M)\frac{N^{(1)}_{\rm g}(M)S(\boldsymbol{r},M)}{[1+\beta(M)N^{(1)}_{\rm g}(M)S(\boldsymbol{r},M)]^2}\,. \tag{71}$$

From the above equations we also notice the useful relation:

$$\overline{\mathcal{G}}^{(2)}_{(1,0)}(\boldsymbol{r})+\overline{\mathcal{G}}^{(1)}_{(2,0)}(\boldsymbol{r}) = \tilde{w}(\boldsymbol{r})\overline{\mathcal{G}}^{(1)}_{(1,1)}(\boldsymbol{r}) = \tilde{w}^2(\boldsymbol{r})\bar{n}_{\rm eff}(\boldsymbol{r})\,.$$

Replacing all these ingredients back into equation (68), after a little algebra we find that $\tilde{w}$ has the solution:

$$\tilde{w}(\boldsymbol{r}) = 1/\left[c+\bar{n}_{\rm eff}(\boldsymbol{r})\right]\,. \tag{72}$$

On putting together equations (67) and (72), back substituting the constant $c = (1+\alpha)/\overline{P}_i$, we arrive at the general solution for the optimal weights:

$$w(\boldsymbol{r},M) = \frac{b(M)}{[1+\beta(M)N^{(1)}_{\rm g}(M)\mathcal{S}(\boldsymbol{r},M)]}\,\frac{1}{\left[(1+\alpha)+\bar{n}_{\rm eff}(\boldsymbol{r})\overline{P}_i\right]}\,. \tag{73}$$

This expression is the central result of this paper. Before inspecting how these weights behave for a specific case, a number of interesting points may be noted. First, if we were able to identify galaxies in a survey whose host halo masses were drawn from some narrow range, then the first factor in equation (73) would be constant and so the weights would revert back to a scheme that structurally resembles FKP, although with a different effective number density. Secondly, we note that there is no natural limit where the above weighting scheme follows that derived by PVP. This dissimilarity emphasises how important an effect modifications of the underlying model assumptions can be on the matter power spectrum estimation and optimization.

Fig. 1 demonstrates how the optimal weights vary as a function of the galaxies host halo mass and redshift. At low redshifts, the galaxy selection $\mathcal{S} \to 1$. For galaxies that are hosted by low-mass haloes, $N^{(1)}_{\rm g}(M) < 1$ and so $w \propto b(M)$. On the other hand, for the high mass clusters $N^{(1)}_{\rm g}(M) \gg 1$, and $w(\boldsymbol{r},M) \propto b(M)/N^{(1)}_{\rm g}(M)$. Hence, we would expect the galaxies in low-mass haloes to be weighted more strongly than those in high-mass haloes, since the bias is rather a slowly evolving function of halo mass. At higher redshift, we would expect that this trend reverses, since $\{\mathcal{S},\bar{n}_{\rm eff}\} \to 0$ and so the weights effectively follow the bias of the host haloes. These trends are exactly what is seen in the figure. For reference, Fig. 1 also compares the optimal weights with the original FKP weight function, given by: $w_{\rm FKP}(\boldsymbol{r}) \propto [1+\bar{n}(\boldsymbol{r})P(k)]^{-1}$.[3] The upturn at large masses in the left-hand panel, is driven by the mass dependence of the ratio $b(M)/N(M)$. For large masses, $b(M)$ is a steep function of mass $b(M) \propto M^{1.5}$ (Seljak & Warren 2004), whereas for most halo occupation distribution models, $N(M) \propto M^1$ (Zehavi et al. 2011). Hence, leading to an upturn for large masses.

5.3 Time evolution of the optimal weights

Before moving on, we briefly discuss the redshift dependence of the optimal weights in equation (73). So far, we have considered that $\bar{n}(M)$, $b(M)$, $\Phi(L|M)$, $N^{(1)}(M)$, $\beta(M)$ and $\xi(\boldsymbol{r})$ are all independent of time (here we will parametrize time evolution through the comoving distance χ). This is approximately correct if the survey volume is sufficiently small so that these functions do not evolve appreciably over the survey. In general, however, they are time dependent. Therefore we would have $\bar{n}(M) \to \bar{n}(M,\chi)$, $b(M) \to b(M,\chi)$, $\Phi(L|M) \to \Phi(L|M,\chi)$, $N^{(1)}(M) \to N^{(1)}(M,\chi)$, $\beta(M) \to \beta(M,\chi)$ and $\xi(\boldsymbol{r}_1-\boldsymbol{r}_2) \to \xi(\boldsymbol{r}_1-\boldsymbol{r}_2,\chi_1,\chi_2) = D(\chi_1)D(\chi_2)\xi(\boldsymbol{r}_1-\boldsymbol{r}_2)$.

In the last equality, we have assumed that the correlation function obeys linear theory, hence the resulting growth factors. Working under this assumption, we redefine the $\overline{\mathcal{G}}$ functions to absorb the growth factors

$$\overline{\mathcal{G}}^{(n)}_{(l,m)}(\boldsymbol{r}) \equiv \int {\rm d}M\bar{n}(M,\chi)\,[D(\chi)b(M,\chi)]^m\, N^{(n)}_{\rm g}(M,\chi)\,\left[w^l(\boldsymbol{r},M)\mathcal{S}(\boldsymbol{r},M)\right]^n \quad . \tag{74}$$

Formally, this is equivalent to redefining the halo bias parameter: $b(M) \to D(\chi)b(M,\chi)$, and we prefer this latter approach. Thus, equation (73) becomes:

$$w(\boldsymbol{r},M) = \frac{D(\chi)b(M,\chi)}{\left[1+\beta(M,\chi)N^{(1)}_{\rm g}(M,\chi)\mathcal{S}(\boldsymbol{r},M)\right]}\,\frac{1}{[(1+\alpha)+\overline{P}_i\,\bar{n}_{\rm eff}(\chi)]}\,, \tag{75}$$

where the new effective number density is

$$\bar{n}_{\rm eff}(\boldsymbol{r}) \equiv \int {\rm d}M\bar{n}(M,\chi)D^2(\chi)b^2(M,\chi)\left[\frac{N^{(1)}_{\rm g}(M,\chi)S(\boldsymbol{r},M)}{1+\beta(M,\chi)N^{(1)}_{\rm g}(M,\chi)S(\boldsymbol{r},M)}\right]. \tag{76}$$

6 INFORMATION CONTENT OF GALAXY CLUSTERING

The ability of a set of band-power estimates of the galaxy power spectrum to constrain the cosmological model can, theoretically, be determined through construction of the Fisher information matrix. Under the assumption that the density field is Gaussianly distributed, one finds that the power spectrum for a given Fourier mode is exponentially distributed about the mean power, and that the band-power estimate is χ^2 distributed (Takahashi et al. 2011). Owing to the central limit theorem, in the limit of a large number of Fourier modes per k-space shell, the power spectrum estimates thus approach the Gaussian distribution. Under the assumption that the power spectrum estimator is Gaussianly distributed, it can be shown that the Fisher matrix has the form (Tegmark 1997; Tegmark, Taylor & Heavens 1997) (but see Abramo 2012):

$$\mathcal{F}_{\alpha\beta} = \frac{1}{2}{\rm Tr}\left[\boldsymbol{C}^{-1}\boldsymbol{C}_{,\alpha}\boldsymbol{C}^{-1}\boldsymbol{C}_{,\beta}\right] + \sum_{i,j}\frac{\partial\overline{P}_i}{\partial\alpha}C^{-1}_{ij}\frac{\partial\overline{P}_j}{\partial\beta} \approx \sum_{i,j}\frac{\partial\overline{P}_i}{\partial\alpha}C^{-1}_{ij}\frac{\partial\overline{P}_j}{\partial\beta}\,,$$

where the approximate equality follows from the fact that the second term on the right-hand side of the first equality dominates over the first term, since it scales directly in proportion with the number of Fourier modes, whereas the first term is independent of the number of modes. In the above, we have made use of the notation $\partial/\partial\alpha \equiv \partial/\partial\theta_\alpha$ to denote partial derivatives with respect to the cosmological parameters θ_α. On taking the covariance matrix to be diagonal, as is the case in equation (51), the above expression for the Fisher matrix becomes:

$$\mathcal{F}_{\alpha\beta} = \sum_{i,j}\frac{\partial\log\overline{P}_i}{\partial\alpha}\overline{P}_i\frac{\delta^K_{ij}}{\sigma^2_P(k_i)}\overline{P}_j\frac{\partial\log\overline{P}_j}{\partial\beta} = \sum_i\frac{\partial\log\overline{P}_i}{\partial\alpha}\frac{\partial\log\overline{P}_i}{\partial\beta}\left(\frac{\mathcal{S}}{\mathcal{N}}\right)^2(k_i)\,. \tag{77}$$

If we now define the effective survey volume through the expression,

$$V_{\rm eff}(k_i) \equiv \frac{2(2\pi)^3}{V_i}\left(\frac{\mathcal{S}}{\mathcal{N}}\right)^2(k_i)\,, \tag{78}$$

and take the continuum limit for the Fourier modes, we find that the Fisher matrix can be expressed as (Tegmark 1997)

$$\mathcal{F}_{\alpha\beta} = \frac{1}{2}\int\frac{{\rm d}^3k}{(2\pi)^3}\frac{\partial\log P(k)}{\partial\alpha}\frac{\partial\log P(k)}{\partial\beta}V_{\rm eff}(k)\,. \tag{79}$$

[3] Note that in order to evaluate the weight functions we took $m_{\rm lim} = 22$ and adopted the CLF model of Yang et al. (2003) to compute $\mathcal{S}(\boldsymbol{r},M)$ and $N^{(1)}_{\rm g}(M)$. For the function $\beta(M)$ we employed the model presented in Cooray & Sheth (2002) derived from semi-analytic galaxies.

Thus in order to determine the information content of the galaxy power spectrum obtained using a general weight function w, we simply need to calculate $V_{\rm eff}[w](k)$ or equivalently $\mathcal{S}/\mathcal{N}[w](k)$. It is clear from equations (51) and (52) that a general expression for the $\mathcal{S}/\mathcal{N}$ is given by:

$$\left(\frac{\mathcal{S}}{\mathcal{N}}\right)^2(k_i) = \frac{V_i}{2(2\pi)^3}\int {\rm d}^3r \left[\overline{\mathcal{G}}^{(1)}_{(1,1)}(\boldsymbol{r})\right]^2 \left\{\int {\rm d}^3r \left(\left[\overline{\mathcal{G}}^{(1)}_{(1,1)}(\boldsymbol{r})\right]^2 + \frac{(1+\alpha)}{\overline{P}_i}\left[\overline{\mathcal{G}}^{(2)}_{(1,0)}(\boldsymbol{r}) + \overline{\mathcal{G}}^{(1)}_{(2,0)}(\boldsymbol{r})\right]\right)^2\right\}^{-1}. \tag{80}$$

In the case of the optimal weights from equation (75), a little algebra leads to the simplified result:

$$\left(\frac{\mathcal{S}}{\mathcal{N}}\right)^2(k_i) = \frac{V_i}{2(2\pi)^3}\int {\rm d}^3r \left[\frac{\overline{P}_i\bar{n}_{\rm eff}(\boldsymbol{r})}{(1+\alpha)+\overline{P}_i\bar{n}_{\rm eff}(\boldsymbol{r})}\right]^2. \tag{81}$$

The above expression will be useful for forecasting how well a future galaxy redshift survey may constrain cosmological parameters, after an optimal power spectrum analysis has been performed.

7 PRACTICAL CHALLENGES IN IMPLEMENTING THE OPTIMAL WEIGHTS

In order to implement the optimal weighting scheme, one requires knowledge of the halo mass function $\bar{n}(M)$, the halo bias function $b(M)$, the conditional probability density $\Phi(L|M)$, the first and second factorial moments of the halo occupation distribution as parametrized by $N^{(1)}(M)$ and $\beta(M)$ and a way to associate each galaxy in the survey to a host halo. A possible route for achieving this is as follows.

(i) Pure halo-dependent quantities: $\bar{n}(M)$ and $b(M)$: these functions can be determined directly from numerical simulations; there also exist a number of accurate semi-analytic fitting functions in the literature (for recent examples see Tinker et al. 2008; Crocce et al. 2010; Watson et al. 2013). However, in order to employ these one needs to specify the underlying cosmological model – we do not consider this too troublesome since it is also required to turn redshifts into distances.

(ii) Galaxy formation-dependent functions: $\Phi(L|M)$, $N^{(1)}_{\rm g}(M)$, $N^{(2)}_{\rm g}(M)$. These require a model of galaxy formation or additional measurements. On adopting a state-of-the-art SAM, these functions can be measured directly (Benson et al. 2000; Cooray & Sheth 2002). They may also be obtained from the data through the CLF approach (Yang et al. 2003; van den Bosch et al. 2013).

(iii) Associating galaxies to groups: this step could be performed through application of standard friends-of-friends group finding algorithms or more sophisticated colour–magnitude–redshift grouping methods (Eke et al. 2004; Koester et al. 2007; Rykoff et al. 2014).

(iv) Determine group halo mass: through the use of good quality mock catalogues, such as can be facilitated through an SAM, one may apply the same grouping algorithms as were used on the real data to the mock data, thus finding the mapping between each group and the most likely halo mass (Eke et al. 2004).

(v) Implement the optimal weighting scheme and measure $P(k)$.

Owing to the fact that the steps enumerated above cannot be performed without error, it is likely that this will introduce additional scatter that we have not accounted for in our optimal estimator. We expect that this scatter will not bias the measurements, but will most likely lead to a reduction in $\mathcal{S}/\mathcal{N}$. We shall leave it as a task for future work to explore how well this method can be implemented in detail.

8 CONCLUSIONS

In this paper, we have developed the theory for the unbiased and optimal estimation of the matter power spectrum from the galaxy power spectrum. Our approach generalizes the original approach of FKP, by taking into account central ideas from the theory of galaxy formation: galaxies form and reside exclusively in dark matter haloes; a given dark matter halo may host many galaxies of various luminosities; galaxies inherit part of their large-scale bias from their host halo.

In Section 2, we described the generic properties of a galaxy redshift survey and presented a new theoretical quantity: the galaxy-halo double-delta expansion. We demonstrated how one may use this expansion of the halo and galaxy fields to answer basic statistical questions concerning the galaxy distribution. In particular, we gave a derivation of the galaxy luminosity function in this framework.

In Section 3, we presented estimators for the galaxy correlation function and power spectrum. It was proved that, in the large-scale and large-survey-volume limits, these were unbiased estimates of the dark matter correlation function and power spectrum. We demonstrated that, similar to FKP, in our scheme the matter power spectrum could be obtained by subtracting an effective shot-noise component followed by the deconvolution of the power spectrum associated with an effective survey window function.

In Section 4, we derived general expressions for the covariance matrix of the weighted galaxy power spectrum, including all non-Gaussian terms arising from the non-linear evolution of matter fluctuations, discreteness effects and finite survey geometry effects. These results generalize the earlier results of Meiksin & White (1999), Scoccimarro et al. (1999) and Smith (2009). In the limits of large scales, large survey volumes and Gaussian fluctuations, the covariance matrix was found to be diagonal.

In Section 5, we found an equation that governs the optimal weights to be applied to galaxies. We found a general solution of the weight equation. Interestingly, the solution did not carry any explicit dependence on galaxy luminosity. Instead the weights were found to be simply a function of two variables: the spatial position within the survey and the mass of the dark matter halo hosting the galaxies.

In Section 6, we presented a new expression for the Fisher information matrix, for a weighted galaxy power spectrum measurement. We also presented a formula for the $\mathcal{S}/\mathcal{N}$ obtained for the optimal weights.

Finally, in Section 7 we outlined the practical steps that would need to be followed if one were to carry out the optimal power spectrum analysis.

In a companion work (Smith & Marian 2015), we explore the $\mathcal{S}/\mathcal{N}$ and cosmological information gains achievable through the optimal weighting scheme. In a future work, we will also explore how well one may implement such a scheme with real data.

ACKNOWLEDGEMENTS

We thank Simon White for useful discussions. RES acknowledges part support from ERC Advanced grant 246797 GALFORMOD. LM thanks MPA for its kind hospitality while part of this work was being performed.

REFERENCES

Abramo L. R., 2012, MNRAS, 420, 2042
Anderson L. et al., 2012, MNRAS, 427, 3435
Anderson L. et al., 2014a, MNRAS, 439, 83
Anderson L. et al., 2014b, MNRAS, 441, 24
Baumgart D. J., Fry J. N., 1991, ApJ, 375, 25
Benson A. J., Cole S., Frenk C. S., Baugh C. M., Lacey C. G., 2000, MNRAS, 311, 793
Bernstein G. M., 1994, ApJ, 424, 569
Blake C. A., Abdalla F. B., Bridle S. L., Rawlings S., 2004, New Astron. Rev., 48, 1063
Blake C. et al., 2011, MNRAS, 415, 2876
Blake C. et al., 2013, MNRAS, 436, 3089
Brown M. J. I., Webster R. L., Boyle B. J., 2000, MNRAS, 317, 782
Cooray A., Sheth R., 2002, Phys. Rep., 372, 1
Crocce M., Fosalba P., Castander F. J., Gaztañaga E., 2010, MNRAS, 403, 1353
Davis M., Geller M. J., 1976, ApJ, 208, 13
Davis M., Peebles P. J. E., 1983, ApJ, 267, 465
Eke V. R. et al., 2004, MNRAS, 348, 866
Feldman H. A., Kaiser N., Peacock J. A., 1994, ApJ, 426, 23 (FKP)
Fisher K. B., Davis M., Strauss M. A., Yahil A., Huchra J. P., 1993, ApJ, 402, 42
Fry J. N., Gaztanaga E., 1993, ApJ, 413, 447
Fry J. N., Peebles P. J. E., 1978, ApJ, 221, 19
Fry J. N., Peebles P. J. E., 1980, ApJ, 238, 785
Groth E. J., Peebles P. J. E., 1977, ApJ, 217, 385
Guzzo L., Strauss M. A., Fisher K. B., Giovanelli R., Haynes M. P., 1997, ApJ, 489, 37
Hamilton A. J. S., 1993, ApJ, 417, 19
Hamilton A. J. S., 1997a, MNRAS, 289, 285
Hamilton A. J. S., 1997b, MNRAS, 289, 295
Hamilton A. J. S., 2000, MNRAS, 312, 257
Hamilton A. J. S., Tegmark M., 2000, MNRAS, 312, 285
Hauser M. G., Peebles P. J. E., 1973, ApJ, 185, 757
Kauffmann G., Colberg J. M., Diaferio A., White S. D. M., 1999, MNRAS, 303, 188
Koester B. P. et al., 2007, ApJ, 660, 239
Landy S. D., Szalay A. S., 1993, ApJ, 412, 64
Laureijs R. et al., 2011, preprint (arXiv:1110.3193)
Levi M. et al., 2013, preprint (arXiv:1308.0847)
Li C., Kauffmann G., Jing Y. P., White S. D. M., Börner G., Cheng F. Z., 2006, MNRAS, 368, 21
Meiksin A., White M., 1999, MNRAS, 308, 1179
Mo H. J., White S. D. M., 1996, MNRAS, 282, 347
Mo H. J., Jing Y. P., White S. D. M., 1997, MNRAS, 284, 189
Norberg P. et al., 2001, MNRAS, 328, 64
Norberg P. et al., 2002, MNRAS, 332, 827
Park C., Vogeley M. S., Geller M. J., Huchra J. P., 1994, ApJ, 431, 569
Peacock J. A., Nicholson D., 1991, MNRAS, 253, 307
Peebles P. J. E., 1973, ApJ, 185, 413
Peebles P. J. E., 1974, ApJS, 28, 37
Peebles P. J. E., 1975, ApJ, 196, 647
Peebles P. J. E., Groth E. J., 1975, ApJ, 196, 1
Peebles P. J. E., Hauser M. G., 1974, ApJS, 28, 19
Percival W. J., Verde L., Peacock J. A., 2004, MNRAS, 347, 645 (PVP)
Rykoff E. S. et al., 2014, ApJ, 785, 104
Scoccimarro R., Zaldarriaga M., Hui L., 1999, ApJ, 527, 1
Seldner M., Peebles P. J. E., 1977, ApJ, 215, 703
Seldner M., Peebles P. J. E., 1978, ApJ, 225, 7
Seldner M., Peebles P. J. E., 1979, ApJ, 227, 30
Seljak U., Warren M. S., 2004, MNRAS, 355, 129
Sheth R. K., Lemson G., 1999, MNRAS, 304, 767

Smith R. E., 2009, MNRAS, 1337
Smith R. E., 2012, MNRAS, 426, 531
Smith R. E., Marian L., 2014, preprint (arXiv:1406.1800) (SM14)
Smith R. E., Marian L., 2015, preprint (arXiv:1507.04365)
Smith R. E., Watts P. I. R., 2005, MNRAS, 360, 203
Smith R. E., Scoccimarro R., Sheth R. K., 2007, PRD, 75, 063512
Springel V. et al., 2005, Nature, 435, 629
Swanson M. E. C., Tegmark M., Blanton M., Zehavi I., 2008, MNRAS, 385, 1635
Takahashi R. et al., 2011, ApJ, 726, 7
Tegmark M., 1997, Phys. Rev. Lett., 79, 3806
Tegmark M., Taylor A. N., Heavens A. F., 1997, ApJ, 480, 22
Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709
van den Bosch F. C., More S., Cacciato M., Mo H., Yang X., 2013, MNRAS, 430, 725
Vogeley M. S., Szalay A. S., 1996, ApJ, 465, 34
Watson W. A., Iliev I. T., D'Aloisio A., Knebe A., Shapiro P. R., Yepes G., 2013, MNRAS, 433, 1230
White S. D. M., Frenk C. S., 1991, ApJ, 379, 52
White S. D. M., Rees M. J., 1978, MNRAS, 183, 341
Yang X., Mo H. J., van den Bosch F. C., 2003, MNRAS, 339, 1057
Zehavi I. et al., 2002, ApJ, 571, 172
Zehavi I. et al., 2005, ApJ, 630, 1
Zehavi I. et al., 2011, ApJ, 736, 59

APPENDIX A: DERIVATION OF THE TWO POINT CORRELATIONS: $\langle N_{\rm g}N'_{\rm g}\rangle$, $\langle N_{\rm g}N'_{\rm s}\rangle$ AND $\langle N_{\rm s}N'_{\rm s}\rangle$

Let us begin by defining the short-hand notation for the correlation: $\langle n_{\rm g}n'_{\rm g}\rangle \equiv \langle n_{\rm g}(\boldsymbol{r}, L, \boldsymbol{x}, M)n_{\rm g}(\boldsymbol{r}', L', \boldsymbol{x}', M')\rangle$. Following the analysis of Section 2, this correlation may be written:

$$\langle n_{\rm g}n'_{\rm g}\rangle = \left\langle \sum_{i,j=1}^{N_{\rm h}} \delta^{\rm D}(\boldsymbol{x}-\boldsymbol{x}_i)\delta^{\rm D}(M-M_i)\delta^{\rm D}(\boldsymbol{x}'-\boldsymbol{x}_j)\delta^{\rm D}(M'-M_j) \right.$$
$$\left. \times \left\langle \sum_{k=1}^{N_{\rm g}(M_i)} \sum_{l=1}^{N_{\rm g}(M_j)} \delta^{\rm D}(\boldsymbol{r}-\boldsymbol{r}_k-\boldsymbol{x}_i)\delta^{\rm D}(L-L_k)\delta^{\rm D}(\boldsymbol{r}'-\boldsymbol{r}_l-\boldsymbol{x}_j)\delta^{\rm D}(L'-L_l)\right\rangle_g \right\rangle_{\rm s}. \tag{A1}$$

If we now split the sums over i and j into two parts, a piece where $i \neq j$ and a piece where $i = j$, then we find

$$\langle n_{\rm g}n'_{\rm g}\rangle = \left\langle \sum_{i\neq j}^{N_{\rm h}} \delta^{\rm D}(\boldsymbol{x}-\boldsymbol{x}_i)\delta^{\rm D}(M-M_i)\delta^{\rm D}(\boldsymbol{x}'-\boldsymbol{x}_j)\delta^{\rm D}(M'-M_j) \right.$$
$$\left. \times \left\langle \sum_{k=1}^{N_{\rm g}(M_i)} \sum_{l=1}^{N_{\rm g}(M_j)} \delta^{\rm D}(\boldsymbol{r}-\boldsymbol{r}_k-\boldsymbol{x}_i)\delta^{\rm D}(L-L_k)\delta^{\rm D}(\boldsymbol{r}'-\boldsymbol{r}_l-\boldsymbol{x}_j)\delta^{\rm D}(L'-L_l)\right\rangle_g \right\rangle_{\rm s}$$
$$+\left\langle \sum_{i=j}^{N_{\rm h}} \delta^{\rm D}(\boldsymbol{x}-\boldsymbol{x}_i)\delta^{\rm D}(M-M_i)\delta^{\rm D}(\boldsymbol{x}'-\boldsymbol{x}_i)\delta^{\rm D}(M'-M_i) \right.$$
$$\left. \times \left\langle \sum_{k,l=1}^{N_{\rm g}(M_i)} \delta^{\rm D}(\boldsymbol{r}-\boldsymbol{r}_k-\boldsymbol{x}_i)\delta^{\rm D}(L-L_k)\delta^{\rm D}(\boldsymbol{r}'-\boldsymbol{r}_l-\boldsymbol{x}_i)\delta^{\rm D}(L'-L_l)\right\rangle_g \right\rangle_{\rm s}. \tag{A2}$$

Consider the terms associated with the $i \neq j$ sum, since we have assumed that the galaxy properties hosted by the ith halo are independent of the galaxy properties in the jth halo, we may write the average of these terms as the product of the two averages. Next, consider the term $i = j$, and notice that we may also separate the sum over k and l into two terms, a term with $k \neq l$ and a term with $k = l$. This leads us to write the following expression:

$$\langle n_{\rm g}n'_{\rm g}\rangle = \left\langle \sum_{i\neq j}^{N_{\rm h}} \delta^{\rm D}(\boldsymbol{x}-\boldsymbol{x}_i)\delta^{\rm D}(M-M_i)\delta^{\rm D}(\boldsymbol{x}'-\boldsymbol{x}_j)\delta^{\rm D}(M'-M_j) \right.$$
$$\left. \times \left\langle \sum_{k=1}^{N_{\rm g}(M_i)} \delta^{\rm D}(\boldsymbol{r}-\boldsymbol{r}_k-\boldsymbol{x}_i)\delta^{\rm D}(L-L_k)\right\rangle_g \left\langle \sum_{l=1}^{N_{\rm g}(M_j)} \delta^{\rm D}(\boldsymbol{r}'-\boldsymbol{r}_l-\boldsymbol{x}_j)\delta^{\rm D}(L'-L_l)\right\rangle_g \right\rangle_{\rm s}$$
$$+\left\langle \sum_{i=j}^{N_{\rm h}} \delta^{\rm D}(\boldsymbol{x}-\boldsymbol{x}_i)\delta^{\rm D}(M-M_i)\delta^{\rm D}(\boldsymbol{x}'-\boldsymbol{x}_i)\delta^{\rm D}(M'-M_i) \right.$$

$$\times\left[\left\langle \sum_{k=1}^{N_{\rm g}(M_i)}\sum_{l\neq k}^{N_{\rm g}(M_i)} \delta^{\rm D}(\boldsymbol{r}-\boldsymbol{r}_k-\boldsymbol{x}_i)\delta^{\rm D}(L-L_k)\delta^{\rm D}(\boldsymbol{r}'-\boldsymbol{r}_l-\boldsymbol{x}_i)\delta^{\rm D}(L'-L_l)\right\rangle_g \right.$$
$$\left.\left. +\left\langle \sum_{k=l}^{N_{\rm g}(M_i)} \delta^{\rm D}(\boldsymbol{r}-\boldsymbol{r}_k-\boldsymbol{x}_i)\delta^{\rm D}(L-L_k)\delta^{\rm D}(\boldsymbol{r}'-\boldsymbol{r}_k-\boldsymbol{x}_i)\delta^{\rm D}(L'-L_k)\right\rangle_g\right]\right\rangle_{\rm s} . \tag{A3}$$

We are now able to compute the expectations over the galaxy populations, and with the help of equation (13) we find

$$\left\langle n_{\rm g}n_{\rm g}'\right\rangle = \left\langle \sum_{i\neq j}^{N_{\rm h}} \delta^{\rm D}(\boldsymbol{x}-\boldsymbol{x}_i)\delta^{\rm D}(M-M_i)\delta^{\rm D}(\boldsymbol{x}'-\boldsymbol{x}_j)\delta^{\rm D}(M'-M_j)\right.$$
$$\left.\times N_{\rm g}^{(1)}(M_i)N_{\rm g}^{(1)}(M_j)U(\boldsymbol{r}-\boldsymbol{x}_i|M_i)U(\boldsymbol{r}'-\boldsymbol{x}_j|M_j)\Phi(L|M_i)\Phi(L'|M_j)\Theta(\boldsymbol{r}|L)\Theta(\boldsymbol{r}'|L')\right\rangle_{\rm s}$$
$$+\left\langle \sum_{i=j}^{N_{\rm h}} \delta^{\rm D}(\boldsymbol{x}-\boldsymbol{x}_i)\delta^{\rm D}(M-M_i)\delta^{\rm D}(\boldsymbol{x}'-\boldsymbol{x}_i)\delta^{\rm D}(M'-M_i)\right.$$
$$\times\left[N_{\rm g}^{(2)}(M_i)U(\boldsymbol{r}-\boldsymbol{x}_i|M_i)U(\boldsymbol{r}'-\boldsymbol{x}_i|M_i)\Phi(L|M_i)\Phi(L'|M_i)\Theta(\boldsymbol{r}|L)\Theta(\boldsymbol{r}'|L')\right.$$
$$\left.\left.+N_{\rm g}^{(1)}(M_i)\Phi(L|M_i)U(\boldsymbol{r}-\boldsymbol{x}_i|M_i)\Theta(\boldsymbol{r}|L)\delta^{\rm D}(L-L')\delta^{\rm D}(\boldsymbol{r}-\boldsymbol{r}')\right]\right\rangle_{\rm s} , \tag{A4}$$

where in the above we have used a short-hand notation for the factorial moments of the galaxy numbers:

$$N_{\rm g}^{(l)}(M) \equiv \left\langle N_{\rm g}(N_{\rm g}-1)\dots(N_{\rm g}-l+1)|M\right\rangle = \sum_{N_{\rm g}=0}^{\infty} P(N_{\rm g}|\lambda(M))N_{\rm g}(N_{\rm g}-1)\dots(N_{\rm g}-l+1) . \tag{A5}$$

Let us now deal with the averages over the dark matter haloes and let us write the first and second terms in equation (A4) as $\langle n_{\rm g}n_{\rm g}'\rangle_{\rm A}$ and $\langle n_{\rm g}n_{\rm g}'\rangle_{\rm B}$. Considering the first term, the expectations may be computed as in equation (18), and we find

$$\left\langle n_{\rm g}n_{\rm g}'\right\rangle_{\rm A} = \sum_{i\neq j}^{N_{\rm h}}\int\prod_{v=1}^{N_{\rm h}}\left\{{\rm d}^3x_v dM_v\right\}\, p(\boldsymbol{x}_1,\dots,\boldsymbol{x}_{N_{\rm h}},M_1,\dots,M_{N_{\rm h}})\delta^{\rm D}(\boldsymbol{x}-\boldsymbol{x}_i)\delta^{\rm D}(M-M_i)\delta^{\rm D}(\boldsymbol{x}'-\boldsymbol{x}_j)\delta^{\rm D}(M'-M_j)$$
$$\times N_{\rm g}^{(1)}(M_i)N_{\rm g}^{(1)}(M_j)U(\boldsymbol{r}-\boldsymbol{x}_i|M_i)U(\boldsymbol{r}'-\boldsymbol{x}_j|M_j)\Phi(L|M_i)\Phi(L'|M_j)\Theta(\boldsymbol{r}|L)\Theta(\boldsymbol{r}'|L')$$
$$= N_{\rm h}(N_{\rm h}-1)p(\boldsymbol{x},\boldsymbol{x}',M,M')N_{\rm g}^{(1)}(M)N_{\rm g}^{(1)}(M')U(\boldsymbol{r}-\boldsymbol{x}|M)U(\boldsymbol{r}'-\boldsymbol{x}'|M')\Phi(L|M)\Phi(L'|M')\Theta(\boldsymbol{r}|L)\Theta(\boldsymbol{r}'|L') . \tag{A6}$$

The joint probability density functions for the halo centres and masses may be expressed in terms of products of their one-point probability density functions (PDFs) and correlation functions. For the case of two points, we have:

$$p(\boldsymbol{x}_1,\boldsymbol{x}_1,M_1,M_2) \equiv p(\boldsymbol{x}_1,M_1)p(\boldsymbol{x}_2,M_2)\left[1+\xi^{\rm c}(\boldsymbol{x}_1,\boldsymbol{x}_2,M_1,M_2)\right] = \frac{\bar{n}(M_1)\bar{n}(M_2)}{{N_{\rm h}}^2}\left[1+\xi^{\rm c}(\boldsymbol{x}_1,\boldsymbol{x}_2,M_1,M_2)\right] . \tag{A7}$$

In addition, if we assume that the cluster density field is some local function of the underlying dark matter density (Fry & Gaztanaga 1993; Mo & White 1996; Mo, Jing & White 1997; Smith, Scoccimarro & Sheth 2007), the cross-correlation function of clusters of masses M_1 and M_2, at leading order, can be written:

$$\xi^{\rm c}(|\boldsymbol{x}_1-\boldsymbol{x}_2|,M_1,M_2) = b(M_1)b(M_2)\xi(|\boldsymbol{x}_1-\boldsymbol{x}_2|) , \tag{A8}$$

where $\xi(r)$ is the correlation of the underlying matter fluctuations. On using this relation in equation (A6) we find

$$\left\langle n_{\rm g}n_{\rm g}'\right\rangle_{\rm A} = \bar{n}(M)\bar{n}(M')\left[1+b(M_1)b(M_2)\xi(|\boldsymbol{x}_1-\boldsymbol{x}_2|)\right]N_{\rm g}^{(1)}(M)N_{\rm g}^{(1)}(M')U(\boldsymbol{r}-\boldsymbol{x}|M)U(\boldsymbol{r}'-\boldsymbol{x}'|M')$$
$$\times\Phi(L|M)\Phi(L'|M')\Theta(\boldsymbol{r}|L)\Theta(\boldsymbol{r}'|L') . \tag{A9}$$

Returning now to the second terms in equation (A4) and following a similar derivation to the first term, we find

$$\left\langle n_{\rm g}n_{\rm g}'\right\rangle_{\rm B} = \bar{n}(M)N_{\rm g}^{(2)}(M)\Phi(L|M)\Phi(L'|M)\Theta(\boldsymbol{r}|L)\Theta(\boldsymbol{r}'|L')U(\boldsymbol{r}-\boldsymbol{x}|M)U(\boldsymbol{r}'-\boldsymbol{x}|M)\delta^{\rm D}(\boldsymbol{x}-\boldsymbol{x}')\delta^{\rm D}(M-M')$$
$$+\bar{n}(M)N_{\rm g}^{(1)}(M)\Phi(L|M)\Theta(\boldsymbol{r}|L)U(\boldsymbol{r}-\boldsymbol{x}|M)\delta^{\rm D}(\boldsymbol{x}-\boldsymbol{x}')\delta^{\rm D}(M-M')\delta^{\rm D}(L-L')\delta^{\rm D}(\boldsymbol{r}-\boldsymbol{r}') . \tag{A10}$$

Following the derivation $\langle n_g n'_g\rangle$ we may now straightforwardly write down the results for the cases of the cross- and auto-correlation of the synthetic galaxy-halo field with the real one:

$$\left\langle n_g n'_s \right\rangle = \alpha^{-1}\bar{n}(M)\bar{n}(M')N_g(M)N_g(M')\Theta(\boldsymbol{r}|L)\Theta(\boldsymbol{r}'|L')U(\boldsymbol{r}-\boldsymbol{x}|M)U(\boldsymbol{r}'-\boldsymbol{x}'|M')\Phi(L|M)\Phi(L'|M')\,; \tag{A11}$$

$$\begin{aligned}\left\langle n_s n'_s \right\rangle &= \alpha^{-2}\bar{n}(M)\bar{n}(M')N_g(M)N_g(M')\Theta(\boldsymbol{r}|L)\Theta(\boldsymbol{r}'|L')U(\boldsymbol{r}-\boldsymbol{x}|M)U(\boldsymbol{r}'-\boldsymbol{x}'|M')\Phi(L|M)\Phi(L'|M')\\
&\quad+\alpha^{-1}\bar{n}(M)N_g^{(2)}(M)\Phi(L|M)\Phi(L'|M)\Theta(\boldsymbol{r}|L)\Theta(\boldsymbol{r}'|L')U(\boldsymbol{r}-\boldsymbol{x}|M)U(\boldsymbol{r}'-\boldsymbol{x}|M)\delta^{D}(\boldsymbol{x}-\boldsymbol{x}')\delta^{D}(M-M')\\
&\quad+N_g^{(1)}(M)\Phi(L|M)\Theta(\boldsymbol{r}|L)U(\boldsymbol{r}-\boldsymbol{x}|M)\delta^{D}(\boldsymbol{x}-\boldsymbol{x}')\delta^{D}(M-M')\delta^{D}(L-L')\delta^{D}(\boldsymbol{r}-\boldsymbol{r}')\,,\end{aligned} \tag{A12}$$

where in the above we have made use of the following short-hand notation: $\langle n_g n'_s\rangle \equiv \langle n_g(\boldsymbol{r}, L, \boldsymbol{x}, M)n_s(\boldsymbol{r}', L', \boldsymbol{x}', M')\rangle$ and $\langle n_s n'_s\rangle \equiv \langle n_s(\boldsymbol{r}, L, \boldsymbol{x}, M)n_s(\boldsymbol{r}', L', \boldsymbol{x}', M')\rangle$.

APPENDIX B: THE GALAXY COVARIANCE MATRIX

B1 Comment on the covariance matrix of the matter power spectrum

Starting with equation (46) and inserting equation (43) we find

$$\mathrm{Cov}\left[\hat{P}(\boldsymbol{k}_1), \hat{P}(\boldsymbol{k}_2)\right] = \mathrm{Cov}\left[|\tilde{\mathcal{F}}_g(\boldsymbol{k}_1)|^2, |\tilde{\mathcal{F}}_g(\boldsymbol{k}_2)|^2\right] - \mathrm{Cov}\left[|\tilde{\mathcal{F}}_g(\boldsymbol{k}_1)|^2, P_{\rm shot}\right] - \mathrm{Cov}\left[|\tilde{\mathcal{F}}_g(\boldsymbol{k}_2)|^2, P_{\rm shot}\right] + \mathrm{Var}[P_{\rm shot}]\,, \tag{B1}$$

where

$$\begin{aligned}\mathrm{Cov}\left[|\tilde{\mathcal{F}}_g(\boldsymbol{k}_1)|^2, |\tilde{\mathcal{F}}_g(\boldsymbol{k}_2)|^2\right] &\equiv \left\langle|\tilde{\mathcal{F}}_g(\boldsymbol{k}_1)|^2|\tilde{\mathcal{F}}_g(\boldsymbol{k}_2)|^2\right\rangle - \left\langle|\tilde{\mathcal{F}}_g(\boldsymbol{k}_1)|^2\right\rangle\left\langle|\tilde{\mathcal{F}}_g(\boldsymbol{k}_2)|^2\right\rangle \;;\\
\mathrm{Cov}\left[|\tilde{\mathcal{F}}_g(\boldsymbol{k}_i)|^2, P_{\rm shot}\right] &\equiv \left\langle|\tilde{\mathcal{F}}_g(\boldsymbol{k}_i)|^2 P_{\rm shot}\right\rangle - \left\langle|\tilde{\mathcal{F}}_g(\boldsymbol{k}_i)|^2\right\rangle\langle P_{\rm shot}\rangle \quad ; \; i \in \{1, 2\}\,;\\
\mathrm{Var}[P_{\rm shot}] &\equiv \left\langle P_{\rm shot}^2\right\rangle - \langle P_{\rm shot}\rangle^2\;.\end{aligned}$$

If we assume that the statistical uncertainties are dominated by $|\tilde{\mathcal{F}}_g(\boldsymbol{k}_1)|^2$ and not $P_{\rm shot}$, then we may approximate the covariance matrix as is written in equation (46).

APPENDIX C: DERIVATION OF THE COVARIANCE MATRIX OF THE $\tilde{\mathcal{F}}_g$ POWER SPECTRUM

To begin, we notice that we may rewrite the covariance matrix of $\left|\mathcal{F}_g(k)\right|^2$, which is given in equation (46) as

$$\mathrm{Cov}\left[|\tilde{\mathcal{F}}_g(\boldsymbol{k}_1)|^2, |\tilde{\mathcal{F}}_g(\boldsymbol{k}_2)|^2\right] = \int d^3k_3 d^3k_4 \delta^D(\boldsymbol{k}_1+\boldsymbol{k}_3)\delta^D(\boldsymbol{k}_2+\boldsymbol{k}_4)\left[\left\langle\mathcal{F}_g(\boldsymbol{k}_1)\ldots\mathcal{F}_g(\boldsymbol{k}_4)\right\rangle - \left\langle\mathcal{F}_g(\boldsymbol{k}_1)\mathcal{F}_g(\boldsymbol{k}_3)\right\rangle\left\langle\mathcal{F}_g(\boldsymbol{k}_2)\mathcal{F}_g(\boldsymbol{k}_4)\right\rangle\right]. \tag{C1}$$

We see that in order to proceed we need the four-point function of the $\mathcal{F}_g(\boldsymbol{k})$ modes. On transforming to real space, this requirement is transformed into the need to determine the four-point correlation:

$$\begin{aligned}&\left\langle\mathcal{F}_g(\boldsymbol{k}_1)\ldots\mathcal{F}_g(\boldsymbol{k}_4)\right\rangle - \left\langle\mathcal{F}_g(\boldsymbol{k}_1)\mathcal{F}_g(\boldsymbol{k}_3)\right\rangle\left\langle\mathcal{F}_g(\boldsymbol{k}_2)\mathcal{F}_g(\boldsymbol{k}_4)\right\rangle\\
&\quad= \int d^3r_1\ldots d^3r_4\left[\left\langle\mathcal{F}_g(\boldsymbol{r}_1)\ldots\mathcal{F}_g(\boldsymbol{r}_4)\right\rangle - \left\langle\mathcal{F}_g(\boldsymbol{r}_1)\mathcal{F}_g(\boldsymbol{r}_3)\right\rangle\left\langle\mathcal{F}_g(\boldsymbol{r}_2)\mathcal{F}_g(\boldsymbol{r}_4)\right\rangle\right]e^{i\boldsymbol{k}_1\cdot\boldsymbol{r}_1+\ldots+\boldsymbol{k}_4\cdot\boldsymbol{r}_4}\end{aligned} \tag{C2}$$

C1 Computing the four-point correlation function of $\mathcal{F}_g(\boldsymbol{r})$

Since the terms $\langle\mathcal{F}_g(\boldsymbol{r}_i)\mathcal{F}_g(\boldsymbol{r}_j)\rangle$ are given by equation (27), we are left with the task of computing the four-point correlation function of the field $\mathcal{F}_g(\boldsymbol{r})$. Using our relation equation (8), this is given by

$$\begin{aligned}\left\langle\mathcal{F}_g(\boldsymbol{r}_1)\ldots\mathcal{F}_g(\boldsymbol{r}_4)\right\rangle &= \frac{1}{A^2}\prod_{i=1}^{4}\left\{\int dL_i d^3x_i dM_i\, w(\boldsymbol{r}_i, L_i, \boldsymbol{x}_i, M_i)\right\}\\
&\quad\times\langle\,[n_g(\boldsymbol{r}_1, L_1, \boldsymbol{x}_1, M_1) - \alpha\, n_s(\boldsymbol{r}_1, L_1\boldsymbol{x}_1, M_1)]\ldots\left[n_g(\boldsymbol{r}_4, L_4, \boldsymbol{x}_4, M_4) - \alpha n_s(\boldsymbol{r}_4, L_4, \boldsymbol{x}_4, M_4)\right]\rangle\\
&= \frac{1}{A^2}\prod_{i=1}^{4}\left\{\int dL_i d^3x_i dM_i\, w(\boldsymbol{r}_i, L_i, \boldsymbol{x}_i, M_i)\right\}\{\,\langle n_{g,1}\ldots n_{g,4}\rangle - \alpha[\,\langle n_{g,1}n_{g,2}n_{g,3}n_{s,4}\rangle + 3\mathrm{cyc}]\\
&\quad+\alpha^2[\,\langle n_{g,1}n_{g,2}n_{s,3}n_{s,4}\rangle + 5\mathrm{perm}] - \alpha^3[\,\langle n_{g,1}n_{s,2}n_{s,3}n_{s,4}\rangle + 3\mathrm{cyc}] + \alpha^4\left\langle n_{s,1}\ldots n_{s,4}\right\}\end{aligned} \tag{C3}$$

with the short-hand notation identical to that used in Appendix A: $n_{\mathrm{g},i} \equiv n_{\mathrm{g}}(\boldsymbol{r}_i, L_i, \boldsymbol{x}_i, M_i)$ and $n_{s,i} \equiv n_{\mathrm{s}}(\boldsymbol{r}_i, L_i, \boldsymbol{x}_i, M_i)$. Focusing on the first term in curly brackets on the right-hand side, and if we insert our galaxy-halo double-delta expansion we find

$$\left\langle n'_{\mathrm{g},1}\ldots n'_{\mathrm{g},4}\right\rangle = \left\langle \sum_{i_1,i_2,i_3,i_4=1}^{N_{\mathrm{h}}} \delta^{\mathrm{D}}(\boldsymbol{x}'_1-\boldsymbol{x}_{i_1})\delta^{\mathrm{D}}(M'_1-M_{i_1})\ldots\delta^{\mathrm{D}}(\boldsymbol{x}'_4-\boldsymbol{x}_{i_4})\delta^{\mathrm{D}}(M'_4-M_{i_4})\,\langle \mathrm{g}\rangle \right\rangle_{\mathrm{h}}, \tag{C4}$$

where we have introduced the short-hand notation

$$\langle \mathrm{g}\rangle \equiv \left\langle \sum_{j_1,j_2,j_3,j_4=1}^{N_{\mathrm{g}}} \delta^{\mathrm{D}}(\boldsymbol{r}_1-\boldsymbol{r}_{j_1})\delta^{\mathrm{D}}(L_1-L_{j_1})\ldots\delta^{\mathrm{D}}(\boldsymbol{r}_4-\boldsymbol{r}_{j_4})\delta^{\mathrm{D}}(L_4-L_{j_4}) \right\rangle_{\mathrm{g}}. \tag{C5}$$

As was done for the case of the two-point function, we may now split the sum over haloes into five types of terms:

$$\left\langle n'_{\mathrm{g},1}\ldots n'_{\mathrm{g},4}\right\rangle = \langle\Gamma_1\rangle + \langle\Gamma_2\rangle + \langle\Gamma_3\rangle + \langle\Gamma_4\rangle + \langle\Gamma_5\rangle, \tag{C6}$$

where the terms Γ_i are defined:

$$\begin{aligned}
\Gamma_1 =& \sum_{i_1\neq i_2\neq i_3\neq i_4} \delta^{\mathrm{D}}(\boldsymbol{x}'_1-\boldsymbol{x}_{i_1})\ldots\delta^{\mathrm{D}}(\boldsymbol{x}'_4-\boldsymbol{x}_{i_4})\delta^{\mathrm{D}}(M'_1-M_{i_1})\ldots\delta^{\mathrm{D}}(M'_4-M_{i_4})\,\langle \mathrm{g}\rangle;\\
\Gamma_2 =& \sum_{i_1\neq i_2\neq i_3=i_4} \delta^{\mathrm{D}}(\boldsymbol{x}'_1-\boldsymbol{x}_{i_1})\delta^{\mathrm{D}}(\boldsymbol{x}'_2-\boldsymbol{x}_{i_2})\delta^{\mathrm{D}}(M'_1-M_{i_1})\delta^{\mathrm{D}}(M'_2-M_{i_2})\prod_{p=3}^{4}\left\{\delta^{\mathrm{D}}(\boldsymbol{x}'_p-\boldsymbol{x}_{i_p})\delta^{\mathrm{D}}(M'_p-M_{i_p})\right\}\langle \mathrm{g}\rangle + 5\text{ perms};\\
\Gamma_3 =& \sum_{i_1=i_2\neq i_3=i_4} \prod_{p=1}^{2}\left\{\delta^{\mathrm{D}}(\boldsymbol{x}'_p-\boldsymbol{x}_{i_p})\delta^{\mathrm{D}}(M'_p-M_{i_p})\right\}\prod_{q=3}^{4}\left\{\delta^{\mathrm{D}}(\boldsymbol{x}'_q-\boldsymbol{x}_{i_q})\delta^{\mathrm{D}}(M'_q-M_{i_q})\right\}\langle \mathrm{g}\rangle + \text{two perms};\\
\Gamma_4 =& \sum_{i_1\neq i_2=i_3=i_4} \delta^{\mathrm{D}}(\boldsymbol{x}'_1-\boldsymbol{x}_{i_1})\delta^{\mathrm{D}}(M'_1-M_{i_1})\prod_{p=2}^{4}\left\{\delta^{\mathrm{D}}(\boldsymbol{x}'_p-\boldsymbol{x}_{i_p})\delta^{\mathrm{D}}(M'_p-M_{i_p})\right\}\langle \mathrm{g}\rangle + \text{three perms};\\
\Gamma_5 =& \sum_{i_1=i_2=i_3=i_4} \prod_{p=1}^{4}\left\{\delta^{\mathrm{D}}(\boldsymbol{x}'_p-\boldsymbol{x}_{i_p})\delta^{\mathrm{D}}(M'_p-M_{i_p})\right\}\langle \mathrm{g}\rangle.
\end{aligned} \tag{C7}$$

Computing $\langle\Gamma_1\rangle$. Integrating over the Dirac delta functions and relabelling primed variables to unprimed, we write:

$$\begin{aligned}
\langle\Gamma_1\rangle =& \sum_{i_1\neq i_2\neq i_3\neq i_4}^{N_{\mathrm{h}}} p(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_4,M_1,\ldots,M_4)\,\langle \mathrm{g}\rangle\\
&= N_{\mathrm{h}}(N_{\mathrm{h}}-1)(N_{\mathrm{h}}-2)(N_{\mathrm{h}}-3)p(\boldsymbol{x}_1,M_1)\ldots p(\boldsymbol{x}_4,M_4)[\,1+\left\{\xi^{\mathrm{c}}_{12}+\xi^{\mathrm{c}}_{13}+\xi^{\mathrm{c}}_{14}+\xi^{\mathrm{c}}_{23}+\xi^{\mathrm{c}}_{24}+\xi^{\mathrm{c}}_{34}\right\}\\
&\quad+\left\{\zeta^{\mathrm{c}}_{123}+\zeta^{\mathrm{c}}_{234}+\zeta^{\mathrm{c}}_{341}+\zeta^{\mathrm{c}}_{412}\right\}+\left\{\xi^{\mathrm{c}}_{12}\xi^{\mathrm{c}}_{34}+\xi^{\mathrm{c}}_{13}\xi^{\mathrm{c}}_{24}+\xi^{\mathrm{c}}_{14}\xi^{\mathrm{c}}_{23}\right\}+\eta^{\mathrm{c}}_{1234}]\,\langle \mathrm{g}\rangle\\
&\approx \bar{n}_1\ldots\bar{n}_4[\,1+\left\{\xi^{\mathrm{c}}_{12}+5\text{ perms}\right\}+\left\{\zeta^{\mathrm{c}}_{123}+3\text{ perms}\right\}+\left\{\xi^{\mathrm{c}}_{12}\xi^{\mathrm{c}}_{34}+2\text{ perms}\right\}+\eta^{\mathrm{c}}_{1234}]\,\langle \mathrm{g}\rangle,
\end{aligned}$$

where in the above we have decomposed the joint four-point PDF into its respective one-point moments and set of correlation functions. In this case, ζ and η denote the connected three- and four-point correlation functions, respectively. Note also that we used the following short-hand notation:

$$\xi^{\mathrm{c}}_{ij} \equiv \xi^{\mathrm{c}}(\boldsymbol{x}_i,\boldsymbol{x}_j,M_i,M_j) \quad;\quad \zeta^{\mathrm{c}}_{ijk} \equiv \zeta(\boldsymbol{x}_i,\boldsymbol{x}_j,\boldsymbol{x}_k,M_i,M_j,M_k) \quad;\quad \eta^{\mathrm{c}}_{ijkl} \equiv \eta^{\mathrm{c}}(\boldsymbol{x}_i,\boldsymbol{x}_j,\boldsymbol{x}_k,\boldsymbol{x}_l,M_i,M_j,M_k,M_l). \tag{C8}$$

Computing $\langle\Gamma_2\rangle$. We denote $\delta^{\mathrm{D}}_{\mathrm{h},ij} \equiv \delta^{\mathrm{D}}(M_i-M_j)\delta^{\mathrm{D}}(\boldsymbol{x}_i-\boldsymbol{x}_j)$ and $\delta^{\mathrm{D}}_{\mathrm{g},ij} \equiv \delta^{\mathrm{D}}(L_i-L_j)\delta^{\mathrm{D}}(\boldsymbol{r}_i-\boldsymbol{r}_j)$. Taking the expectations and integrating over the delta functions we find

$$\begin{aligned}
\langle\Gamma_2\rangle =& \sum_{i_1\neq i_2\neq i_3=i_4}^{N_{\mathrm{h}}} p(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3,M_1,M_2,M_3)\delta^{\mathrm{D}}_{\mathrm{h},34}\,\langle g\rangle + \sum_{i_1\neq i_2=i_3\neq i_4}^{N_{\mathrm{h}}} p(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_4,M_1,M_2,M_4)\delta^{\mathrm{D}}_{\mathrm{h},23}\,\langle g\rangle\\
&+\sum_{i_1=i_2\neq i_3\neq i_4}^{N_{\mathrm{h}}} p(\boldsymbol{x}_1,\boldsymbol{x}_3,\boldsymbol{x}_4,M_1,M_3,M_4)\delta^{\mathrm{D}}_{\mathrm{h},12}\,\langle g\rangle + \sum_{i_1=i_3\neq i_2\neq i_4}^{N_{\mathrm{h}}} p(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_4,M_1,M_2,M_4)\delta^{\mathrm{D}}_{\mathrm{h},13}\,\langle g\rangle\\
&+\sum_{i_1=i_4\neq i_2\neq i_3}^{N_{\mathrm{h}}} p(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3,M_1,M_2,M_3)\delta^{\mathrm{D}}_{\mathrm{h},14}\,\langle g\rangle + \sum_{i_2=i_4\neq i_1\neq i_3}^{N_{\mathrm{h}}} p(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3,M_1,M_2,M_3)\delta^{\mathrm{D}}_{\mathrm{h},24}\,\langle g\rangle\\
&= N_{\mathrm{h}}(N_{\mathrm{h}}-1)(N_{\mathrm{h}}-2)[\,p(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3,M_1,M_2,M_3)\,\langle g\rangle\,\delta^{\mathrm{D}}_{\mathrm{h},34} + p(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_4,M_1,M_2,M_4)\,\langle g\rangle\,\delta^{\mathrm{D}}_{\mathrm{h},23}
\end{aligned}$$

$$+ p(\boldsymbol{x}_1, \boldsymbol{x}_3, \boldsymbol{x}_4, M_1, M_3, M_4)\,\langle g\rangle\,\delta^{\rm D}_{{\rm h},12} + p(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_4, M_1, M_2, M_4)\,\langle g\rangle\,\delta^{\rm D}_{{\rm h},13} + p(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, M_1, M_2, M_3)\,\langle g\rangle\,\delta^{\rm D}_{{\rm h},14}$$
$$+ p(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, M_1, M_2, M_3)\,\langle g\rangle\,\delta^{\rm D}_{{\rm h},24}\,]$$
$$\approx \bar{n}_1\bar{n}_3\bar{n}_4\left[1+\xi^{\rm c}_{13}+\xi^{\rm c}_{14}+\xi^{\rm c}_{34}+\zeta^{\rm c}_{134}\right]\langle g\rangle\,\delta^{\rm D}_{{\rm h},12} + \bar{n}_1\bar{n}_2\bar{n}_4\left[1+\xi^{\rm c}_{12}+\xi^{\rm c}_{24}+\xi^{\rm c}_{24}+\zeta^{\rm c}_{124}\right]\langle g\rangle\,\delta^{\rm D}_{{\rm h},23}$$
$$+ \bar{n}_1\bar{n}_2\bar{n}_3\left[1+\xi^{\rm c}_{12}+\xi^{\rm c}_{13}+\xi^{\rm c}_{23}+\zeta^{\rm c}_{123}\right]\langle g\rangle\,\delta^{\rm D}_{{\rm h},14} + \bar{n}_1\bar{n}_2\bar{n}_4\left[1+\xi^{\rm c}_{12}+\xi^{\rm c}_{14}+\xi^{\rm c}_{24}+\zeta^{\rm c}_{124}\right]\langle g\rangle\,\delta^{\rm D}_{{\rm h},23}$$
$$+ \bar{n}_1\bar{n}_2\bar{n}_3\left[1+\xi^{\rm c}_{12}+\xi^{\rm c}_{23}+\xi^{\rm c}_{31}+\zeta^{\rm c}_{123}\right]\langle g\rangle\,\delta^{\rm D}_{{\rm h},24} + \bar{n}_1\bar{n}_2\bar{n}_3\left[1+\xi^{\rm c}_{12}+\xi^{\rm c}_{23}+\xi^{\rm c}_{31}+\zeta^{\rm c}_{123}\right]\langle g\rangle\,\delta^{\rm D}_{{\rm h},34}.$$

Computing $\langle\Gamma_3\rangle$. Again, on taking the expectations and integrating over the delta functions we find

$$\langle\Gamma_3\rangle = \sum_{i_1=i_2\neq i_3=i_4} p(\boldsymbol{x}_1, \boldsymbol{x}_3, M_1, M_3)\,\langle g\rangle\,\delta^{\rm D}_{{\rm h},12}\,\delta^{\rm D}_{{\rm h},34} + \sum_{i_1=i_3\neq i_2=i_4} p(\boldsymbol{x}_1, \boldsymbol{x}_2, M_1, M_2)\,\langle g\rangle\,\delta^{\rm D}_{{\rm h},13}\,\delta^{\rm D}_{{\rm h},24} + \sum_{i_1=i_4\neq i_2=i_3} p(\boldsymbol{x}_1, \boldsymbol{x}_2, M_1, M_2)\,\langle g\rangle\,\delta^{\rm D}_{{\rm h},14}\,\delta^{\rm D}_{{\rm h},23}$$
$$= N_{\rm h}(N_{\rm h}-1)\left[p(\boldsymbol{x}_1, \boldsymbol{x}_3, M_1, M_3)\,\langle g\rangle\,\delta^{\rm D}_{{\rm h},12}\,\delta^{\rm D}_{{\rm h},34} + p(\boldsymbol{x}_1, \boldsymbol{x}_2, M_1, M_2)\,\langle g\rangle\,\delta^{\rm D}_{{\rm h},13}\,\delta^{\rm D}_{{\rm h},24} + p(\boldsymbol{x}_1, \boldsymbol{x}_2, M_1, M_2)\,\langle g\rangle\,\delta^{\rm D}_{{\rm h},14}\,\delta^{\rm D}_{{\rm h},23}\right]$$
$$= \bar{n}_1\bar{n}_3\left[1+\xi^{\rm g}_{13}\right]\langle g\rangle\,\delta^{\rm D}_{{\rm h},12}\,\delta^{\rm D}_{{\rm h},34} + \bar{n}_1\bar{n}_2\left[1+\xi^{\rm g}_{12}\right]\langle g\rangle\left[\,\delta^{\rm D}_{{\rm h},13}\,\delta^{\rm D}_{{\rm h},24}+\delta^{\rm D}_{{\rm h},14}\,\delta^{\rm D}_{{\rm h},23}\right].$$

Computing $\langle\Gamma_4\rangle$. Again, on taking the expectations and integrating over the delta functions we find

$$\langle\Gamma_4\rangle = \sum_{i_1=i_2=i_3\neq i_4} p(\boldsymbol{x}_1, \boldsymbol{x}_4, M_1, M_4)\,\langle g\rangle\,\delta^{\rm D}_{{\rm h},12}\,\delta^{\rm D}_{{\rm h},13} + \sum_{i_1=i_2=i_4\neq i_3} p(\boldsymbol{x}_1, \boldsymbol{x}_3, M_1, M_3)\,\langle g\rangle\,\delta^{\rm D}_{{\rm h},12}\,\delta^{\rm D}_{{\rm h},14}$$
$$+ \sum_{i_1=i_3=i_4\neq i_2} p(\boldsymbol{x}_1, \boldsymbol{x}_2, M_1, M_2)\,\langle g\rangle\,\delta^{\rm D}_{{\rm h},13}\,\delta^{\rm D}_{{\rm h},14} + \sum_{i_2=i_3=i_4\neq i_1} p(\boldsymbol{x}_1, \boldsymbol{x}_2, M_1, M_2)\,\langle g\rangle\,\delta^{\rm D}_{{\rm h},23}\,\delta^{\rm D}_{{\rm h},24}$$
$$= N(N-1)[\; p(\boldsymbol{x}_1, \boldsymbol{x}_4, M_1, M_4)\,\langle g\rangle\,\delta^{\rm D}_{{\rm h},12}\,\delta^{\rm D}_{{\rm h},13} + p(\boldsymbol{x}_1, \boldsymbol{x}_3, M_1, M_3)\,\langle g\rangle\,\delta^{\rm D}_{{\rm h},12}\,\delta^{\rm D}_{{\rm h},14}$$
$$+ p(\boldsymbol{x}_1, \boldsymbol{x}_2, M_1, M_2)\,\langle g\rangle\,\delta^{\rm D}_{{\rm h},13}\,\delta^{\rm D}_{{\rm h},14} + p(\boldsymbol{x}_1, \boldsymbol{x}_2, M_1, M_2)\,\langle g\rangle\,\delta^{\rm D}_{{\rm h},23}\,\delta^{\rm D}_{{\rm h},24}\;]$$
$$= \bar{n}_1\bar{n}_4\left[1+\xi^{\rm g}_{14}\right]\langle g\rangle\,\delta^{\rm D}_{{\rm h},12}\,\delta^{\rm D}_{{\rm h},13} + \bar{n}_1\bar{n}_3\left[1+\xi^{\rm g}_{13}\right]\langle g\rangle\,\delta^{\rm D}_{{\rm h},12}\,\delta^{\rm D}_{{\rm h},14} + \bar{n}_1\bar{n}_2\left[1+\xi^{\rm g}_{12}\right]\langle g\rangle\left[\delta^{\rm D}_{{\rm h},13}\,\delta^{\rm D}_{{\rm h},14}+\delta^{\rm D}_{{\rm h},23}\,\delta^{\rm D}_{{\rm h},24}\right]$$

Computing $\langle\Gamma_5\rangle$. Again, on taking the expectations and integrating over the delta functions we find

$$\langle\Gamma_5\rangle = \sum_{i_1=i_2=i_3=i_4} p(\boldsymbol{x}_1, M_1)\,\langle g\rangle\,\delta^{\rm D}_{{\rm h},12}\,\delta^{\rm D}_{{\rm h},13}\,\delta^{\rm D}_{{\rm h},14} = \bar{n}_1\,\langle g\rangle\,\delta^{\rm D}_{{\rm h},12}\,\delta^{\rm D}_{{\rm h},13}\,\delta^{\rm D}_{{\rm h},14}$$

Collecting the terms $\langle\Gamma_1\rangle$, $\langle\Gamma_2\rangle$, $\langle\Gamma_3\rangle$, $\langle\Gamma_4\rangle$ and $\langle\Gamma_5\rangle$, we write:

$$\left\langle n_{{\rm g},1}n_{{\rm g},2}n_{{\rm g},3}n_{{\rm g},4}\right\rangle = \bar{n}_1\bar{n}_2\bar{n}_3\bar{n}_4\Bigg\{\left[1+\xi^{\rm c}_{12}+\xi^{\rm c}_{13}+\xi^{\rm c}_{14}+\xi^{\rm c}_{23}+\xi^{\rm c}_{24}+\xi^{\rm c}_{34}+\zeta^{\rm c}_{123}+\zeta^{\rm c}_{124}+\zeta^{\rm c}_{134}+\zeta^{\rm c}_{234}+\xi^{\rm c}_{12}\xi^{\rm c}_{34}+\xi^{\rm c}_{13}\xi^{\rm c}_{24}\right.$$
$$\left.+\;\xi^{\rm c}_{14}\xi^{\rm c}_{23}+\eta^{\rm c}_{1234}\right]\langle g\rangle + \left[1+\xi^{\rm c}_{13}+\xi^{\rm c}_{14}+\xi^{\rm c}_{34}+\zeta^{\rm c}_{134}\right]\frac{\delta^{\rm D}_{{\rm h},12}}{\bar{n}_2}\langle g\rangle + \left[1+\xi^{\rm c}_{12}+\xi^{\rm c}_{24}+\xi^{\rm c}_{41}+\zeta^{\rm c}_{124}\right]\frac{\delta^{\rm D}_{{\rm h},23}}{\bar{n}_3}\langle g\rangle$$
$$+\left[1+\xi^{\rm c}_{12}+\xi^{\rm c}_{13}+\xi^{\rm c}_{23}+\zeta^{\rm c}_{123}\right]\frac{\delta^{\rm D}_{{\rm h},14}}{\bar{n}_4}\langle g\rangle + \left[1+\xi^{\rm c}_{12}+\xi^{\rm c}_{14}+\xi^{\rm c}_{24}+\zeta^{\rm c}_{124}\right]\frac{\delta^{\rm D}_{{\rm h},13}}{\bar{n}_3}\langle g\rangle$$
$$+\left[1+\xi^{\rm c}_{12}+\xi^{\rm c}_{23}+\xi^{\rm c}_{31}+\zeta^{\rm c}_{123}\right]\frac{\delta^{\rm D}_{{\rm h},24}}{\bar{n}_4}\langle g\rangle + \left[1+\xi^{\rm c}_{12}+\xi^{\rm c}_{23}+\xi^{\rm c}_{31}+\zeta^{\rm c}_{123}\right]\frac{\delta^{\rm D}_{{\rm h},34}}{\bar{n}_4}\langle g\rangle$$
$$+\left[1+\xi^{\rm c}_{13}\right]\frac{\delta^{\rm D}_{{\rm h},12}\delta^{\rm D}_{{\rm h},34}}{\bar{n}_2\bar{n}_4}\langle g\rangle + \left[1+\xi^{\rm c}_{12}\right]\frac{\delta^{\rm D}_{{\rm h},13}\delta^{\rm D}_{{\rm h},24}}{\bar{n}_3\bar{n}_4}\langle g\rangle + \left[1+\xi^{\rm c}_{12}\right]\frac{\delta^{\rm D}_{{\rm h},14}\delta^{\rm D}_{{\rm h},23}}{\bar{n}_3\bar{n}_4}\langle g\rangle + \left[1+\xi^{\rm c}_{14}\right]\frac{\delta^{\rm D}_{{\rm h},12}\delta^{\rm D}_{{\rm h},13}}{\bar{n}_2\bar{n}_3}\langle g\rangle$$
$$+\left[1+\xi^{\rm c}_{13}\right]\frac{\delta^{\rm D}_{{\rm h},12}\delta^{\rm D}_{{\rm h},14}}{\bar{n}_2\bar{n}_4}\langle g\rangle + \left[1+\xi^{\rm c}_{12}\right]\frac{\delta^{\rm D}_{{\rm h},13}\delta^{\rm D}_{{\rm h},14}}{\bar{n}_3\bar{n}_4}\langle g\rangle + \left[1+\xi^{\rm c}_{12}\right]\frac{\delta^{\rm D}_{{\rm h},23}\delta^{\rm D}_{{\rm h},24}}{\bar{n}_3\bar{n}_4}\langle g\rangle + \frac{\delta^{\rm D}_{{\rm h},12}\delta^{\rm D}_{{\rm h},13}\delta^{\rm D}_{{\rm h},14}}{\bar{n}_2\bar{n}_3\bar{n}_4}\langle g\rangle\Bigg\}.$$

Based on the above expression, we are now in a position to immediately write down the other four-point function cases required to compute equation (C3):

$$\left\langle n_{{\rm g},1}n_{{\rm g},2}n_{{\rm g},3}n_{s,4}\right\rangle = \alpha^{-1}\bar{n}_1\bar{n}_2\bar{n}_3\bar{n}_4\Bigg\{\left(1+\xi^{\rm c}_{12}+\xi^{\rm c}_{13}+\xi^{\rm c}_{23}+\zeta^{\rm c}_{123}\right)\langle g\rangle + \left(1+\xi^{\rm c}_{13}\right)\frac{\delta^{\rm D}_{{\rm h},12}}{\bar{n}_2}\langle g\rangle$$
$$+\left(1+\xi^{\rm c}_{12}\right)\left[\frac{\delta^{\rm D}_{{\rm h},23}}{\bar{n}_3}\langle g\rangle + \frac{\delta^{\rm D}_{{\rm h},13}}{\bar{n}_3}\langle g\rangle\right] + \frac{\delta^{\rm D}_{{\rm h},12}\delta^{\rm D}_{{\rm h},13}}{\bar{n}_2\bar{n}_3}\langle g\rangle\Bigg\}; \qquad \text{(C9)}$$

$$\left\langle n_{\mathrm{g},1} n_{\mathrm{g},2} n_{s,3} n_{s,4} \right\rangle = \alpha^{-2} \bar{n}_1 \bar{n}_2 \bar{n}_3 \bar{n}_4 \left\{ \left(1+\xi^{\mathrm{c}}_{12}\right) \langle g \rangle + \frac{\delta^{\mathrm{D}}_{\mathrm{h},12}}{\bar{n}_2} \langle g \rangle + \alpha \left(1+\xi^{\mathrm{c}}_{12}\right) \frac{\delta^{\mathrm{D}}_{\mathrm{h},34}}{\bar{n}_4} \langle g \rangle + \alpha \frac{\delta^{\mathrm{D}}_{\mathrm{h},12}\,\delta^{\mathrm{D}}_{\mathrm{h},34}}{\bar{n}_2 \bar{n}_4} \langle g \rangle \right\}; \tag{C10}$$

$$\left\langle n_{\mathrm{g},1} n_{s,2} n_{s,3} n_{s,4} \right\rangle = \alpha^{-3} \bar{n}_1 \bar{n}_2 \bar{n}_3 \bar{n}_4 \left\{ \langle g \rangle + \alpha \left[\frac{\delta^{\mathrm{D}}_{\mathrm{h},23}}{\bar{n}_3} + \frac{\delta^{\mathrm{D}}_{\mathrm{h},24}}{\bar{n}_4} + \frac{\delta^{\mathrm{D}}_{\mathrm{h},34}}{\bar{n}_4} \right] \langle g \rangle + \alpha \frac{\delta^{\mathrm{D}}_{\mathrm{h},23} \delta^{\mathrm{D}}_{\mathrm{h},24}}{\bar{n}_3 \bar{n}_4} \langle g \rangle \right\}; \tag{C11}$$

$$\begin{aligned} \left\langle n_{s,1} n_{s,2} n_{s,3} n_{s,4} \right\rangle = \alpha^{-4} \bar{n}_1 \bar{n}_2 \bar{n}_3 \bar{n}_4 \Bigg\{ & \langle g \rangle + \alpha \left[\frac{\delta^{\mathrm{D}}_{\mathrm{h},12}}{\bar{n}_2} \langle g \rangle + \frac{\delta^{\mathrm{D}}_{\mathrm{h},23}}{\bar{n}_3} \langle g \rangle + \frac{\delta^{\mathrm{D}}_{\mathrm{h},13}}{\bar{n}_3} \langle g \rangle + \frac{\delta^{\mathrm{D}}_{\mathrm{h},14}}{\bar{n}_4} \langle g \rangle + \frac{\delta^{\mathrm{D}}_{\mathrm{h},24}}{\bar{n}_4} \langle g \rangle + \frac{\delta^{\mathrm{D}}_{\mathrm{h},34}}{\bar{n}_4} \langle g \rangle \right] \\ & + \alpha^2 \left[\frac{\delta^{\mathrm{D}}_{\mathrm{h},12} \delta^{\mathrm{D}}_{\mathrm{h},34}}{\bar{n}_2 \bar{n}_4} \langle g \rangle + \frac{\delta^{\mathrm{D}}_{\mathrm{h},13} \delta^{\mathrm{D}}_{\mathrm{h},24}}{\bar{n}_3 \bar{n}_4} \langle g \rangle + \frac{\delta^{\mathrm{D}}_{\mathrm{h},23} \delta^{\mathrm{D}}_{\mathrm{h},14}}{\bar{n}_3 \bar{n}_4} \langle g \rangle + \frac{\delta^{\mathrm{D}}_{\mathrm{h},12} \delta^{\mathrm{D}}_{\mathrm{h},13}}{\bar{n}_2 \bar{n}_3} \langle g \rangle + \frac{\delta^{\mathrm{D}}_{\mathrm{h},12} \delta^{\mathrm{D}}_{\mathrm{h},14}}{\bar{n}_2 \bar{n}_4} \langle g \rangle \right. \\ & \left. + \frac{\delta^{\mathrm{D}}_{\mathrm{h},13} \delta^{\mathrm{D}}_{\mathrm{h},14}}{\bar{n}_3 \bar{n}_4} \langle g \rangle + \frac{\delta^{\mathrm{D}}_{\mathrm{h},23} \delta^{\mathrm{D}}_{\mathrm{h},24}}{\bar{n}_3 \bar{n}_4} \langle g \rangle \right] + \alpha^3 \frac{\delta^{\mathrm{D}}_{\mathrm{h},12} \delta^{\mathrm{D}}_{\mathrm{h},13} \delta^{\mathrm{D}}_{\mathrm{h},14}}{\bar{n}_2 \bar{n}_3 \bar{n}_4} \langle g \rangle \Bigg\}. \end{aligned} \tag{C12}$$

In order to compute the four-point correlation function of $\mathcal{F}_{\mathrm{g}}$, we need to compute the sum of four terms of equation (C11), with permuted location of g and s. This is given by

$$\begin{aligned} L_1 &\equiv \left\langle n_{\mathrm{g},1} n_{s,2} n_{s,3} n_{s,4} \right\rangle + \left\langle n_{s,1} n_{\mathrm{g},2} n_{s,3} n_{s,4} \right\rangle + \left\langle n_{s,1} n_{s,2} n_{\mathrm{g},3} n_{s,4} \right\rangle + \left\langle n_{s,1} n_{s,2} n_{s,3} n_{\mathrm{g},4} \right\rangle \\ &= \bar{n}_1 \bar{n}_2 \bar{n}_3 \bar{n}_4 \Bigg\{ 4\alpha^{-3} \langle g \rangle + 2\alpha^{-2} \left[\frac{\delta^{\mathrm{D}}_{\mathrm{h},23}}{\bar{n}_3} \langle g \rangle + \frac{\delta^{\mathrm{D}}_{\mathrm{h},24}}{\bar{n}_4} \langle g \rangle + \frac{\delta^{\mathrm{D}}_{\mathrm{h},34}}{\bar{n}_4} \langle g \rangle + \frac{\delta^{\mathrm{D}}_{\mathrm{h},13}}{\bar{n}_3} \langle g \rangle + \frac{\delta^{\mathrm{D}}_{\mathrm{h},14}}{\bar{n}_4} \langle g \rangle + \frac{\delta^{\mathrm{D}}_{\mathrm{h},12}}{\bar{n}_2} \langle g \rangle \right] \\ &\quad + \alpha^{-1} \left[\frac{\delta^{\mathrm{D}}_{\mathrm{h},23}}{\bar{n}_3} \frac{\delta^{\mathrm{D}}_{\mathrm{h},24}}{\bar{n}_4} \langle g \rangle + \frac{\delta^{\mathrm{D}}_{\mathrm{h},13}}{\bar{n}_3} \frac{\delta^{\mathrm{D}}_{\mathrm{h},14}}{\bar{n}_4} \langle g \rangle + \frac{\delta^{\mathrm{D}}_{\mathrm{h},12}}{\bar{n}_2} \frac{\delta^{\mathrm{D}}_{\mathrm{h},14}}{\bar{n}_4} \langle g \rangle + \frac{\delta^{\mathrm{D}}_{\mathrm{h},12}}{\bar{n}_2} \frac{\delta^{\mathrm{D}}_{\mathrm{h},13}}{\bar{n}_3} \langle g \rangle \right] \Bigg\}. \end{aligned}$$

We also need to compute the sum of six terms of equation (C10), with permuted location of g and s. This is given by

$$\begin{aligned} L_2 &\equiv \left\langle n_{\mathrm{g},1} n_{\mathrm{g},2} n_{s,3} n_{s,4} \right\rangle + \text{five perms} \\ &= \bar{n}_1 \bar{n}_2 \bar{n}_3 \bar{n}_4 \alpha^{-2} \Bigg\{ \left(6 + \xi^{\mathrm{c}}_{12} + \xi^{\mathrm{c}}_{13} + \xi^{\mathrm{c}}_{14} + \xi^{\mathrm{c}}_{23} + \xi^{\mathrm{c}}_{24} + \xi^{\mathrm{c}}_{34}\right) \langle g \rangle + \frac{\delta^{\mathrm{D}}_{\mathrm{h},12}}{\bar{n}_2} \langle g \rangle + \left[\frac{\delta^{\mathrm{D}}_{\mathrm{h},13}}{\bar{n}_3} + \frac{\delta^{\mathrm{D}}_{\mathrm{h},23}}{\bar{n}_3} \right] \langle g \rangle \\ &+ \left[\frac{\delta^{\mathrm{D}}_{\mathrm{h},14}}{\bar{n}_4} + \frac{\delta^{\mathrm{D}}_{\mathrm{h},24}}{\bar{n}_4} + \frac{\delta^{\mathrm{D}}_{\mathrm{h},34}}{\bar{n}_4} \right] \langle g \rangle + \alpha \Bigg(\left[\left(1+\xi^{\mathrm{c}}_{12}\right) \frac{\delta^{\mathrm{D}}_{\mathrm{h},34}}{\bar{n}_4} + \left(1+\xi^{\mathrm{c}}_{13}\right) \frac{\delta^{\mathrm{D}}_{\mathrm{h},24}}{\bar{n}_4} + \left(1+\xi^{\mathrm{c}}_{23}\right) \frac{\delta^{\mathrm{D}}_{\mathrm{h},14}}{\bar{n}_4} \right] \langle g \rangle \\ &+ \left[\left(1+\xi^{\mathrm{c}}_{14}\right) \frac{\delta^{\mathrm{D}}_{\mathrm{h},23}}{\bar{n}_3} \langle g \rangle + \left(1+\xi^{\mathrm{c}}_{24}\right) \frac{\delta^{\mathrm{D}}_{\mathrm{h},13}}{\bar{n}_3} \right] \langle g \rangle + \left(1+\xi^{\mathrm{c}}_{34}\right) \frac{\delta^{\mathrm{D}}_{\mathrm{h},12}}{\bar{n}_2} \langle g \rangle + 2 \frac{\delta^{\mathrm{D}}_{\mathrm{h},12}}{\bar{n}_2} \frac{\delta^{\mathrm{D}}_{\mathrm{h},34}}{\bar{n}_4} \langle g \rangle + 2 \left[\frac{\delta^{\mathrm{D}}_{\mathrm{h},13}}{\bar{n}_3} \frac{\delta^{\mathrm{D}}_{\mathrm{h},24}}{\bar{n}_4} + \frac{\delta^{\mathrm{D}}_{\mathrm{h},23}}{\bar{n}_3} \frac{\delta^{\mathrm{D}}_{\mathrm{h},14}}{\bar{n}_4} \right] \langle g \rangle \Bigg) \Bigg\}. \end{aligned}$$

In addition, we need to compute the sum of four terms of equation (C9), again where the locations of g and s are permuted. This sum can be written:

$$\begin{aligned} L_3 &\equiv \langle n_{\mathrm{g},1} n_{\mathrm{g},2} n_{\mathrm{g},3} n_{s,4} \rangle + \langle n_{\mathrm{g},1} n_{\mathrm{g},2} n_{s,3} n_{\mathrm{g},4} \rangle + \langle n_{\mathrm{g},1} n_{s,2} n_{\mathrm{g},3} n_{\mathrm{g},4} \rangle + \langle n_{s,1} n_{\mathrm{g},2} n_{\mathrm{g},3} n_{\mathrm{g},4} \rangle \\ &= \alpha^{-1} \bar{n}_1 \bar{n}_2 \bar{n}_3 \bar{n}_4 \Bigg\{ \left(4 + 2\xi^{\mathrm{c}}_{12} + 2\xi^{\mathrm{c}}_{13} + 2\xi^{\mathrm{c}}_{23} + 2\xi^{\mathrm{c}}_{14} + 2\xi^{\mathrm{c}}_{24} + 2\xi^{\mathrm{c}}_{34} + \zeta^{\mathrm{c}}_{123} + \zeta^{\mathrm{c}}_{124} + \zeta^{\mathrm{c}}_{134} + \zeta^{\mathrm{c}}_{234}\right) \langle g \rangle \\ &\quad + \left(1+\xi^{\mathrm{c}}_{13}\right) \frac{\delta^{\mathrm{D}}_{\mathrm{h},12}}{\bar{n}_2} \langle g \rangle + \left(1+\xi^{\mathrm{c}}_{12}\right) \frac{\delta^{\mathrm{D}}_{\mathrm{h},23}}{\bar{n}_3} \langle g \rangle + \left(1+\xi^{\mathrm{c}}_{12}\right) \frac{\delta^{\mathrm{D}}_{\mathrm{h},31}}{\bar{n}_3} \langle g \rangle + \left(1+\xi^{\mathrm{c}}_{14}\right) \frac{\delta^{\mathrm{D}}_{\mathrm{h},12}}{\bar{n}_2} \langle g \rangle + \left(1+\xi^{\mathrm{c}}_{12}\right) \frac{\delta^{\mathrm{D}}_{\mathrm{h},24}}{\bar{n}_4} \langle g \rangle \\ &\quad + \left(1+\xi^{\mathrm{c}}_{12}\right) \frac{\delta^{\mathrm{D}}_{\mathrm{h},41}}{\bar{n}_4} \langle g \rangle + \left(1+\xi^{\mathrm{c}}_{14}\right) \frac{\delta^{\mathrm{D}}_{\mathrm{h},13}}{\bar{n}_3} \langle g \rangle + \left(1+\xi^{\mathrm{c}}_{13}\right) \frac{\delta^{\mathrm{D}}_{\mathrm{h},14}}{\bar{n}_4} \langle g \rangle + \left(1+\xi^{\mathrm{c}}_{13}\right) \frac{\delta^{\mathrm{D}}_{\mathrm{h},34}}{\bar{n}_4} \langle g \rangle + \left(1+\xi^{\mathrm{c}}_{24}\right) \frac{\delta^{\mathrm{D}}_{\mathrm{h},23}}{\bar{n}_3} \langle g \rangle \\ &\quad + \left(1+\xi^{\mathrm{c}}_{23}\right) \frac{\delta^{\mathrm{D}}_{\mathrm{h},24}}{\bar{n}_4} \langle g \rangle + \left(1+\xi^{\mathrm{c}}_{23}\right) \frac{\delta^{\mathrm{D}}_{\mathrm{h},34}}{\bar{n}_4} \langle g \rangle + \frac{\delta^{\mathrm{D}}_{\mathrm{h},12}}{\bar{n}_2} \frac{\delta^{\mathrm{D}}_{\mathrm{h},13}}{\bar{n}_3} \langle g \rangle + \frac{\delta^{\mathrm{D}}_{\mathrm{h},12}}{\bar{n}_2} \frac{\delta^{\mathrm{D}}_{\mathrm{h},14}}{\bar{n}_4} \langle g \rangle + \frac{\delta^{\mathrm{D}}_{\mathrm{h},13}}{\bar{n}_3} \frac{\delta^{\mathrm{D}}_{\mathrm{h},14}}{\bar{n}_4} \langle g \rangle + \frac{\delta^{\mathrm{D}}_{\mathrm{h},23}}{\bar{n}_3} \frac{\delta^{\mathrm{D}}_{\mathrm{h},24}}{\bar{n}_4} \langle g \rangle \Bigg\}. \end{aligned}$$

Collecting the terms L_1, L_2 and L_3 along with $\langle n_{\mathrm{g},1} n_{\mathrm{g},2} n_{\mathrm{g},3} n_{\mathrm{g},4} \rangle$ and $\langle n_{s,1} n_{s,2} n_{s,3} n_{s,4} \rangle$ and inserting them into equation (C3), and after some algebra we arrive at the following arrangement:

$$\begin{aligned} \left\langle F_{\mathrm{g},1} \ldots F_{\mathrm{g},4} \right\rangle = \frac{1}{A^2} \prod_{i=1}^{4} \left\{ \int \mathrm{d}L_i \mathrm{d}^3 x_i \, dM_i \, \bar{n}_i w_i \right\} \Bigg\{ & \eta^{\mathrm{c}}_{1234} \langle g \rangle + \left[\xi^{\mathrm{c}}_{12} + \frac{(1+\alpha)}{\bar{n}_2} \delta^{\mathrm{D}}_{\mathrm{h},12} \right] \left[\xi^{\mathrm{c}}_{34} + \frac{(1+\alpha)}{\bar{n}_4} \delta^{\mathrm{D}}_{\mathrm{h},34} \right] \langle g \rangle \\ & + \left[\xi^{\mathrm{c}}_{13} + \frac{(1+\alpha)}{\bar{n}_3} \delta^{\mathrm{D}}_{\mathrm{h},13} \right] \left[\xi^{\mathrm{c}}_{24} + \frac{(1+\alpha)}{\bar{n}_4} \delta^{\mathrm{D}}_{\mathrm{h},24} \right] \langle g \rangle + \left[\xi^{\mathrm{c}}_{14} + \frac{(1+\alpha)}{\bar{n}_4} \delta^{\mathrm{D}}_{\mathrm{h},14} \right] \left[\xi^{\mathrm{c}}_{23} + \frac{(1+\alpha)}{\bar{n}_3} \delta^{\mathrm{D}}_{\mathrm{h},23} \right] \langle g \rangle \end{aligned}$$

$$+\zeta^{\rm c}_{134}\frac{\delta^{\rm D}_{\rm h,12}}{\bar{n}_2}\langle g\rangle+\zeta^{\rm c}_{124}\frac{\delta^{\rm D}_{\rm h,23}}{\bar{n}_3}\langle g\rangle+\zeta^{\rm c}_{123}\frac{\delta^{\rm D}_{\rm h,14}}{\bar{n}_4}\langle g\rangle+\zeta^{\rm c}_{124}\frac{\delta^{\rm D}_{\rm h,13}}{\bar{n}_3}\langle g\rangle+\zeta^{\rm c}_{123}\frac{\delta^{\rm D}_{\rm h,24}}{\bar{n}_4}\langle g\rangle+\zeta^{\rm c}_{123}\frac{\delta^{\rm D}_{\rm h,34}}{\bar{n}_4}\langle g\rangle$$
$$+\xi^{\rm c}_{13}\frac{\delta^{\rm D}_{\rm h,12}\delta^{\rm D}_{\rm h,34}}{\bar{n}_2\bar{n}_4}\langle g\rangle+\xi^{\rm c}_{12}\frac{\delta^{\rm D}_{\rm h,13}\delta^{\rm D}_{\rm h,24}}{\bar{n}_3\bar{n}_4}\langle g\rangle+\xi^{\rm c}_{12}\frac{\delta^{\rm D}_{\rm h,14}\delta^{\rm D}_{\rm h,23}}{\bar{n}_3\bar{n}_4}\langle g\rangle+\xi^{\rm c}_{14}\frac{\delta^{\rm D}_{\rm h,12}\delta^{\rm D}_{\rm h,13}}{\bar{n}_2\bar{n}_3}\langle g\rangle+\xi^{\rm c}_{13}\frac{\delta^{\rm D}_{\rm h,12}\delta^{\rm D}_{\rm h,14}}{\bar{n}_2\bar{n}_4}\langle g\rangle$$
$$+\xi^{\rm c}_{12}\frac{\delta^{\rm D}_{\rm h,13}\delta^{\rm D}_{\rm h,14}}{\bar{n}_3\bar{n}_4}\langle g\rangle+\xi^{\rm c}_{12}\frac{\delta^{\rm D}_{\rm h,23}\delta^{\rm D}_{\rm h,24}}{\bar{n}_3\bar{n}_4}\langle g\rangle+\frac{(1+\alpha^3)\delta^{\rm D}_{\rm h,12}\delta^{\rm D}_{\rm h,13}\delta^{\rm D}_{\rm h,14}}{\bar{n}_2\bar{n}_3\bar{n}_4}\langle g\rangle\Bigg\},$$

where in the above we used the short-hand notation $F_{{\rm g},i}\equiv\mathcal{F}_{\rm g}(\boldsymbol{r}_i)$. In order to compute the covariance, we also require the second term in equation (C2). On repeatedly using equation (27) we find that this can be written:

$$\left\langle F_{\rm g,1}F_{\rm g,3}\right\rangle\left\langle F_{\rm g,2}F_{\rm g,4}\right\rangle=\prod_{i=1}^{4}\left\{\int dL_i{\rm d}^3x_i{\rm d}M_i\,\bar{n}_iw_i\right\}\left[\xi^{\rm c}_{13}\langle g\rangle+\frac{(1+\alpha)\delta^{\rm D}_{\rm h,13}}{\bar{n}_3}\langle g\rangle\right]\left[\xi^{\rm c}_{24}\langle g\rangle+\frac{(1+\alpha)\delta^{\rm D}_{\rm h,24}}{\bar{n}_4}\langle g\rangle\right].$$

Joining the last two equations, we write the covariance as:

$$\left\langle\mathcal{F}_{\rm g}(\boldsymbol{r}_1)\mathcal{F}_{\rm g}(\boldsymbol{r}_2)\mathcal{F}_{\rm g}(\boldsymbol{r}_3)\mathcal{F}_{\rm g}(\boldsymbol{r}_4)\right\rangle-\left\langle\mathcal{F}_{\rm g}(\boldsymbol{r}_1)\mathcal{F}_{\rm g}(\boldsymbol{r}_3)\right\rangle\left\langle\mathcal{F}_{\rm g}(\boldsymbol{r}_2)\mathcal{F}_{\rm g}(\boldsymbol{r}_4)\right\rangle=\frac{1}{A^2}\prod_{i=1}^{4}\left\{\int{\rm d}L_i{\rm d}^3x_i{\rm d}M_i\,\bar{n}_iw_i\right\}\Bigg\{\eta^{\rm c}_{1234}\langle g\rangle$$
$$+\left[\xi^{\rm c}_{12}+\frac{(1+\alpha)}{\bar{n}_2}\delta^{\rm D}_{\rm h,12}\right]\left[\xi^{\rm c}_{34}+\frac{(1+\alpha)}{\bar{n}_4}\delta^{\rm D}_{\rm h,34}\right]\langle g\rangle+\left[\xi^{\rm c}_{14}+\frac{(1+\alpha)}{\bar{n}_4}\delta^{\rm D}_{\rm h,14}\right]\left[\xi^{\rm c}_{23}+\frac{(1+\alpha)}{\bar{n}_3}\delta^{\rm D}_{\rm h,23}\right]\langle g\rangle+\zeta^{\rm c}_{134}\frac{\delta^{\rm D}_{\rm h,12}}{\bar{n}_2}\langle g\rangle$$
$$+\zeta^{\rm c}_{124}\frac{\delta^{\rm D}_{\rm h,23}}{\bar{n}_3}\langle g\rangle+\zeta^{\rm c}_{123}\frac{\delta^{\rm D}_{\rm h,14}}{\bar{n}_4}\langle g\rangle+\zeta^{\rm c}_{124}\frac{\delta^{\rm D}_{\rm h,13}}{\bar{n}_3}\langle g\rangle+\zeta^{\rm c}_{123}\frac{\delta^{\rm D}_{\rm h,24}}{\bar{n}_4}\langle g\rangle+\zeta^{\rm c}_{123}\frac{\delta^{\rm D}_{\rm h,34}}{\bar{n}_4}\langle g\rangle+\xi^{\rm c}_{13}\frac{\delta^{\rm D}_{\rm h,12}\delta^{\rm D}_{\rm h,34}}{\bar{n}_2\bar{n}_4}\langle g\rangle+\xi^{\rm c}_{12}\frac{\delta^{\rm D}_{\rm h,13}\delta^{\rm D}_{\rm h,24}}{\bar{n}_3\bar{n}_4}\langle g\rangle$$
$$+\xi^{\rm c}_{12}\frac{\delta^{\rm D}_{\rm h,14}\delta^{\rm D}_{\rm h,23}}{\bar{n}_3\bar{n}_4}\langle g\rangle+\xi^{\rm c}_{14}\frac{\delta^{\rm D}_{\rm h,12}\delta^{\rm D}_{\rm h,13}}{\bar{n}_2\bar{n}_3}\langle g\rangle+\xi^{\rm c}_{13}\frac{\delta^{\rm D}_{\rm h,12}\delta^{\rm D}_{\rm h,14}}{\bar{n}_2\bar{n}_4}\langle g\rangle+\xi^{\rm c}_{12}\frac{\delta^{\rm D}_{\rm h,13}\delta^{\rm D}_{\rm h,14}}{\bar{n}_3\bar{n}_4}\langle g\rangle+\xi^{\rm c}_{12}\frac{\delta^{\rm D}_{\rm h,23}\delta^{\rm D}_{\rm h,24}}{\bar{n}_3\bar{n}_4}\langle g\rangle+\frac{(1+\alpha^3)\delta^{\rm D}_{\rm h,12}\delta^{\rm D}_{\rm h,13}\delta^{\rm D}_{\rm h,14}}{\bar{n}_2\bar{n}_3\bar{n}_4}\langle g\rangle\Bigg\}.$$
$$\text{(C13)}$$

We now make the assumption that the fluctuations are close to Gaussian; hence, we take $\eta=\zeta=0$.

$$\left\langle\mathcal{F}_{\rm g}(\boldsymbol{r}_1)\mathcal{F}_{\rm g}(\boldsymbol{r}_2)\mathcal{F}_{\rm g}(\boldsymbol{r}_3)\mathcal{F}_{\rm g}(\boldsymbol{r}_4)\right\rangle-\left\langle\mathcal{F}_{\rm g}(\boldsymbol{r}_1)\mathcal{F}_{\rm g}(\boldsymbol{r}_3)\right\rangle\left\langle\mathcal{F}_{\rm g}(\boldsymbol{r}_2)\mathcal{F}_{\rm g}(\boldsymbol{r}_4)\right\rangle=\frac{1}{A^2}\prod_{i=1}^{4}\left\{\int{\rm d}L_i{\rm d}^3x_i{\rm d}M_i\,\bar{n}_iw_i\right\}\Bigg\{\frac{(1+\alpha^3)\delta^{\rm D}_{\rm h,12}\delta^{\rm D}_{\rm h,13}\delta^{\rm D}_{\rm h,14}}{\bar{n}_2\bar{n}_3\bar{n}_4}\langle g\rangle$$
$$\left[\xi^{\rm c}_{12}+\frac{(1+\alpha)}{\bar{n}_2}\delta^{\rm D}_{\rm h,12}\right]\left[\xi^{\rm c}_{34}+\frac{(1+\alpha)}{\bar{n}_4}\delta^{\rm D}_{\rm h,34}\right]\langle g\rangle+\left[\xi^{\rm c}_{14}+\frac{(1+\alpha)}{\bar{n}_4}\delta^{\rm D}_{\rm h,14}\right]\left[\xi^{\rm c}_{23}+\frac{(1+\alpha)}{\bar{n}_3}\delta^{\rm D}_{\rm h,23}\right]\langle g\rangle+\xi^{\rm c}_{14}\frac{\delta^{\rm D}_{\rm h,12}\delta^{\rm D}_{\rm h,13}}{\bar{n}_2\bar{n}_3}\langle g\rangle$$
$$+\xi^{\rm c}_{13}\left[\frac{\delta^{\rm D}_{\rm h,12}\delta^{\rm D}_{\rm h,14}}{\bar{n}_2\bar{n}_4}\langle g\rangle+\frac{\delta^{\rm D}_{\rm h,12}\delta^{\rm D}_{\rm h,34}}{\bar{n}_2\bar{n}_4}\langle g\rangle\right]+\xi^{\rm c}_{12}\left[\frac{\delta^{\rm D}_{\rm h,13}\delta^{\rm D}_{\rm h,14}}{\bar{n}_3\bar{n}_4}\langle g\rangle+\frac{\delta^{\rm D}_{\rm h,13}\delta^{\rm D}_{\rm h,24}}{\bar{n}_3\bar{n}_4}\langle g\rangle+\frac{\delta^{\rm D}_{\rm h,23}\delta^{\rm D}_{\rm h,14}}{\bar{n}_3\bar{n}_4}\langle g\rangle+\frac{\delta^{\rm D}_{\rm h,23}\delta^{\rm D}_{\rm h,24}}{\bar{n}_3\bar{n}_4}\langle g\rangle\right]\Bigg\}.\quad\text{(C14)}$$

On taking the limit that $\bar{n}_{\rm g}V_\mu\gg 1$, the first and last three terms will be subdominant (Smith 2009). Using the linear bias model from equation (A8), we write

$$\left\langle\mathcal{F}_{\rm g}(\boldsymbol{r}_1)\mathcal{F}_{\rm g}(\boldsymbol{r}_2)\mathcal{F}_{\rm g}(\boldsymbol{r}_3)\mathcal{F}_{\rm g}(\boldsymbol{r}_4)\right\rangle-\left\langle\mathcal{F}_{\rm g}(\boldsymbol{r}_1)\mathcal{F}_{\rm g}(\boldsymbol{r}_3)\right\rangle\left\langle\mathcal{F}_{\rm g}(\boldsymbol{r}_2)\mathcal{F}_{\rm g}(\boldsymbol{r}_4)\right\rangle=\frac{1}{A^2}\prod_{i=1}^{4}\left\{\int{\rm d}L_i{\rm d}^3x_i{\rm d}M_i\,\bar{n}_iw_i\right\}$$
$$\times\Bigg\{\left[b_1b_2\xi_{12}+\frac{(1+\alpha)}{\bar{n}_2}\delta^{\rm D}_{\rm h,12}\right]\left[b_3b_4\xi_{34}+\frac{(1+\alpha)}{\bar{n}_4}\delta^{\rm D}_{\rm h,34}\right]\langle g\rangle+\left[b_1b_4\xi_{14}+\frac{(1+\alpha)}{\bar{n}_4}\delta^{\rm D}_{\rm h,14}\right]\left[b_2b_3\xi_{23}+\frac{(1+\alpha)}{\bar{n}_3}\delta^{\rm D}_{\rm h,23}\right]\langle g\rangle\Bigg\}.\quad\text{(C15)}$$

C2 Averaging over the galaxy distributions

Let us now return to the evaluation of the expectation values for the galaxy population. Consider again equation (C15) and let us look in particular at the terms $\langle g\rangle$ and any pre-multiplying Dirac delta functions. For compactness, we shall use the following definition:

$$f(p|q)\equiv\theta(\boldsymbol{r}_p|L_p)\Phi(L_p|M_q)U(\boldsymbol{r}_p-\boldsymbol{x}_q|M_q)\quad\text{(C16)}$$

We find that there are three types of terms forming the individual $\langle g\rangle$ factors:

$$\langle g\rangle \rightarrow \prod_{p=1}^{4}\left\{N^{(1)}_{\mathrm{g},p} f(p|p)\right\} ;$$

$$\delta^{\mathrm{D}}_{\mathrm{h},14}\langle g\rangle \rightarrow \prod_{p=1}^{3}\left\{N^{(1)}_{\mathrm{g},p} f(p|p)\right\}\left[\frac{N^{(2)}_{\mathrm{g},1}}{N^{(1)}_{\mathrm{g},1}} f(4|1)+\delta^{\mathrm{D}}_{\mathrm{g},14}\right]\delta^{\mathrm{D}}_{\mathrm{h},14} ;$$

$$\delta^{\mathrm{D}}_{\mathrm{h},12}\delta^{\mathrm{D}}_{\mathrm{h},34}\langle g\rangle \rightarrow \prod_{p\in\{1,3\}}\left\{N^{(1)}_{\mathrm{g},p} f(p|p)\right\}\left[\frac{N^{(2)}_{\mathrm{g},1}}{N^{(1)}_{\mathrm{g},1}} f(2|1)+\delta^{\mathrm{D}}_{\mathrm{g},12}\right]\left[\frac{N^{(2)}_{\mathrm{g},3}}{N^{(1)}_{\mathrm{g},3}} f(4|3)+\delta^{\mathrm{D}}_{\mathrm{g},34}\right]\delta^{\mathrm{D}}_{\mathrm{h},12}\delta^{\mathrm{D}}_{\mathrm{h},34}. \quad \text{(C17)}$$

The rest of the $\langle g\rangle$ factors can be worked out in a similar way. Embedding them into equation (C15), and using equations (C1) and (C2), we arrive at the expression for the covariance of the power spectrum estimator in the Gaussian approximation:

$$\begin{aligned}\mathrm{Cov}\left[|\mathcal{F}_{\mathrm{g}}(\boldsymbol{k}_1)|^2, |\mathcal{F}_{\mathrm{g}}(\boldsymbol{k}_2)|^2\right] &= \left|\int \frac{\mathrm{d}^3 q}{(2\pi)^3} P(\boldsymbol{q}) G_{(1,1)}(\boldsymbol{k}_1,\boldsymbol{q}) G_{(1,1)}(\boldsymbol{k}_2,-\boldsymbol{q}) + (1+\alpha)\left[G_{(2,0)}(\boldsymbol{k}_1+\boldsymbol{k}_2,\boldsymbol{0}) + G(\boldsymbol{k}_1,\boldsymbol{k}_2)\right]\right|^2\\ &+\left|\int \frac{\mathrm{d}^3 q}{(2\pi)^3} P(\boldsymbol{q}) G_{(1,1)}(\boldsymbol{k}_1,\boldsymbol{q}) G_{(1,1)}(-\boldsymbol{k}_2,-\boldsymbol{q}) + (1+\alpha)\left[G_{(2,0)}(\boldsymbol{k}_1-\boldsymbol{k}_2,\boldsymbol{0}) + G(\boldsymbol{k}_1,-\boldsymbol{k}_2)\right]\right|^2, \end{aligned} \quad \text{(C18)}$$

where we have defined two more functions:

$$G_{(l,m)}(\boldsymbol{k},\boldsymbol{q}) \equiv \int \mathrm{d}M \bar{n}(M) b^m(M) N^{(1)}_{\mathrm{g}}(M) \tilde{\mathcal{W}}^U_{(l)}(\boldsymbol{k},\boldsymbol{q},M) ;$$

$$G(\boldsymbol{k}_1,\boldsymbol{k}_2) \equiv \int \mathrm{d}M \bar{n}(M) N^{(2)}_{\mathrm{g}}(M) \int \frac{\mathrm{d}^3 q}{(2\pi)^3} \tilde{\mathcal{W}}^U_{(1)}(\boldsymbol{k}_1,\boldsymbol{q},M) \tilde{\mathcal{W}}^U_{(1)}(\boldsymbol{k}_2,-\boldsymbol{q},M). \quad \text{(C19)}$$

C3 The covariance matrix in the large-scale limit

In the large-scale limit, the profiles of the galaxies behave like Dirac delta functions, e.g. $U(\boldsymbol{r}-\boldsymbol{x}|M) \rightarrow \delta^{\mathrm{D}}(\boldsymbol{r}-\boldsymbol{x})$. It is straightforward to show that in this limit the above-defined functions become

$$G_{(l,m)}(\boldsymbol{k},\boldsymbol{q}) \rightarrow \tilde{\mathcal{G}}^{(1)}_{(l,m)}(\boldsymbol{k}+\boldsymbol{q}) \quad ; \quad G(\boldsymbol{k}_1,\boldsymbol{k}_2) \rightarrow \tilde{\mathcal{G}}^{(2)}_{(1,0)}(\boldsymbol{k}_1+\boldsymbol{k}_2),$$

where $\tilde{\mathcal{G}}^{(n)}_{(l,m)}$ are the Fourier transforms of the functions defined by equation (36). With these changes, equation (C18) can be expressed as

$$\begin{aligned}\mathrm{Cov}\left[|\mathcal{F}_{\mathrm{g}}(\boldsymbol{k}_1)|^2, |\mathcal{F}_{\mathrm{g}}(\boldsymbol{k}_2)|^2\right] &= \left|\int \frac{\mathrm{d}^3 q}{(2\pi)^3} P(\boldsymbol{q}) \tilde{\mathcal{G}}^{(1)}_{(1,1)}(\boldsymbol{k}_1+\boldsymbol{q}) \tilde{\mathcal{G}}^{(1)}_{(1,1)}(\boldsymbol{k}_2-\boldsymbol{q}) + (1+\alpha)\left[\tilde{\mathcal{G}}^{(1)}_{(2,0)}(\boldsymbol{k}_1+\boldsymbol{k}_2) + \tilde{\mathcal{G}}^{(2)}_{(1,0)}(\boldsymbol{k}_1+\boldsymbol{k}_2)\right]\right|^2\\ &+\left|\int \frac{\mathrm{d}^3 q}{(2\pi)^3} P(\boldsymbol{q}) \tilde{\mathcal{G}}^{(1)}_{(1,1)}(\boldsymbol{k}_1+\boldsymbol{q}) \tilde{\mathcal{G}}^{(1)}_{(1,1)}(-\boldsymbol{k}_2-\boldsymbol{q}) + (1+\alpha)\left[\tilde{\mathcal{G}}^{(1)}_{(2,0)}(\boldsymbol{k}_1-\boldsymbol{k}_2) + \tilde{\mathcal{G}}^{(2)}_{(1,0)}(\boldsymbol{k}_1-\boldsymbol{k}_2)\right]\right|^2, \end{aligned}$$

which is exactly equation (47). This concludes our proof of it.

APPENDIX D: SHELL AVERAGING THE COVARIANCE MATRIX OF THE $\tilde{\mathcal{F}}_{\mathrm{g}}$ POWER SPECTRUM

A reasonable approximation when computing the shell-averaged power is that if the shells are narrow compared to the scale over which the power spectrum varies, one can factor the latter out of the integrals in equation (50), writing:

$$\begin{aligned}\mathrm{Cov}\left[|\tilde{\mathcal{F}}_{\mathrm{g}}(k_i)|^2, |\tilde{\mathcal{F}}_{\mathrm{g}}(k_j)|^2\right] &= 2\overline{P}^2(k_i)\int_{V_i}\frac{\mathrm{d}^3 k_1}{V_i}\int_{V_j}\frac{\mathrm{d}^3 k_2}{V_j}\tilde{\mathcal{Q}}^{(1,1)}_{(1,1|1,1)}(\boldsymbol{k}_1+\boldsymbol{k}_2)\tilde{\mathcal{Q}}^{(1,1)}_{(1,1|1,1)}(-\boldsymbol{k}_1-\boldsymbol{k}_2)\\ &+4(1+\alpha)\overline{P}(k_i)\int_{V_i}\frac{\mathrm{d}^3 k_1}{V_i}\int_{V_j}\frac{\mathrm{d}^3 k_2}{V_j}\tilde{\mathcal{Q}}^{(1,1)}_{(1,1|1,1)}(\boldsymbol{k}_1+\boldsymbol{k}_2)\tilde{\mathcal{Q}}^{(2)}_{(1|0)}(-\boldsymbol{k}_1-\boldsymbol{k}_2)\\ &+4(1+\alpha)\overline{P}(k_i)\int_{V_i}\frac{\mathrm{d}^3 k_1}{V_i}\int_{V_j}\frac{\mathrm{d}^3 k_2}{V_j}\tilde{\mathcal{Q}}^{(1,1)}_{(1,1|1,1)}(\boldsymbol{k}_1+\boldsymbol{k}_2)\tilde{\mathcal{Q}}^{(1)}_{(2|0)}(-\boldsymbol{k}_1-\boldsymbol{k}_2)\\ &+4(1+\alpha)^2\int_{V_i}\frac{\mathrm{d}^3 k_1}{V_i}\int_{V_j}\frac{\mathrm{d}^3 k_2}{V_j}\tilde{\mathcal{Q}}^{(2)}_{(1|0)}(\boldsymbol{k}_1+\boldsymbol{k}_2)\tilde{\mathcal{Q}}^{(1)}_{(2|0)}(-\boldsymbol{k}_1-\boldsymbol{k}_2)\\ &+2(1+\alpha)^2\int_{V_i}\frac{\mathrm{d}^3 k_1}{V_i}\int_{V_j}\frac{\mathrm{d}^3 k_2}{V_j}\tilde{\mathcal{Q}}^{(2)}_{(1|0)}(\boldsymbol{k}_1+\boldsymbol{k}_2)\tilde{\mathcal{Q}}^{(2)}_{(1|0)}(-\boldsymbol{k}_1-\boldsymbol{k}_2)\\ &+2(1+\alpha)^2\int_{V_i}\frac{\mathrm{d}^3 k_1}{V_i}\int_{V_j}\frac{\mathrm{d}^3 k_2}{V_j}\tilde{\mathcal{Q}}^{(1)}_{(2|0)}(\boldsymbol{k}_1+\boldsymbol{k}_2)\tilde{\mathcal{Q}}^{(1)}_{(2|0)}(-\boldsymbol{k}_1-\boldsymbol{k}_2). \end{aligned} \quad \text{(D1)}$$

Our task now is to solve the integrals forming the terms of equation (D1). In general, this integrals have the form:

$$\begin{aligned}
&\int_{V_i}\frac{d^3k_1}{V_i}\int_{V_j}\frac{d^3k_2}{V_j}\tilde{\mathcal{Q}}^{(n_1,n_2)}_{(l_1,l_2|m_1,m_2)}(\boldsymbol{k}_1+\boldsymbol{k}_2)\tilde{\mathcal{Q}}^{(n_1',n_2')}_{(l_1',l_2'|m_1',m_2')}(-\boldsymbol{k}_1-\boldsymbol{k}_2)\\
&\quad=\int_{V_i}\frac{d^3k_1}{V_i}\int_{V_j}\frac{d^3k_2}{V_j}\int d^3r_1d^3r_2\,\mathcal{Q}^{(n_1,n_2)}_{(l_1,l_2|m_1,m_2)}(\boldsymbol{r}_1)\mathcal{Q}^{(n_1',n_2')}_{(l_1',l_2'|m_1',m_2')}(\boldsymbol{r}_2)e^{i(\boldsymbol{k}_1+\boldsymbol{k}_2)\cdot(\boldsymbol{r}_1-\boldsymbol{r}_2)}\\
&\quad=\int d^3r_1d^3r_2\,\mathcal{Q}^{(n_1,n_2)}_{(l_1,l_2|m_1,m_2)}(\boldsymbol{r}_1)\mathcal{Q}^{(n_1',n_2')}_{(l_1',l_2'|m_1',m_2')}(\boldsymbol{r}_2)\int_{V_i}\frac{d^3k_1}{V_i}e^{i\boldsymbol{k}_1\cdot(\boldsymbol{r}_1-\boldsymbol{r}_2)}\int_{V_j}\frac{d^3k_2}{V_j}e^{i\boldsymbol{k}_2\cdot(\boldsymbol{r}_1-\boldsymbol{r}_2)}\\
&\quad=\int d^3r_1d^3r_2\,\mathcal{Q}^{(n_1,n_2)}_{(l_1,l_2|m_1,m_2)}(\boldsymbol{r}_1)\mathcal{Q}^{(n_1',n_2')}_{(l_1',l_2'|m_1',m_2')}(\boldsymbol{r}_2)\overline{j_0}(k_i|\boldsymbol{r}_1-\boldsymbol{r}_2|)\overline{j_0}(k_j|\boldsymbol{r}_1-\boldsymbol{r}_2|)\\
&\quad=\int d^3r_{21}\,\overline{j_0}(k_ir_{21})\overline{j_0}(k_jr_{21})\Sigma^{(n_1,n_2)(n_1',n_2')}_{(l_1,l_2|m_1,m_2)(l_1',l_2'|m_1',m_2')}(r_{21})
\end{aligned}\tag{D2}$$

In the above, we have defined the shell average of the spherical Bessel function as

$$\overline{j_0}(k_ir)\equiv\frac{1}{V_i}\int_{k_i-\Delta k/2}^{k_i+\Delta k/2}dk_1k_1^2 4\pi j_0(k_1r)\,.\tag{D3}$$

To obtain the last line of equation (D2), we made a change of variables $\boldsymbol{r}_{21}=\boldsymbol{r}_2-\boldsymbol{r}_1$, and defined the correlation function of the weighted survey window function to be

$$\Sigma^{(n_1,n_2)(n_1',n_2')}_{(l_1,l_2|m_1,m_2)(l_1',l_2'|m_1',m_2')}(r_{21})\equiv\int\frac{d^2\hat{\boldsymbol{r}}_{21}}{4\pi}\int d^3r_1\,\mathcal{Q}^{(n_1,n_2)}_{(l_1,l_2|m_1,m_2)}(\boldsymbol{r}_1)\mathcal{Q}^{(n_1',n_2')}_{(l_1',l_2'|m_1',m_2')}(\boldsymbol{r}_{21}+\boldsymbol{r}_1)\,.\tag{D4}$$

In the limit that the survey volume is large, the weighted survey window correlation function is very slowly varying over nearly all length scales of interest, and so can be approximated by its value at zero lag. Using the orthogonality relation of the Bessel functions, $\int_0^\infty drr^2j_\alpha(ur)j_\alpha(vr)=(\pi/2u^2)\delta^D(u-v)$ we write:

$$\int_{V_i}\frac{d^3k_1}{V_i}\int_{V_j}\frac{d^3k_2}{V_j}\tilde{\mathcal{Q}}^{(n_1,n_2)}_{(l_1,l_2|m_1,m_2)}(\boldsymbol{k}_1+\boldsymbol{k}_2)\tilde{\mathcal{Q}}^{(n_1',n_2')}_{(l_1',l_2'|m_1',m_2')}(-\boldsymbol{k}_1-\boldsymbol{k}_2)\approx\frac{(2\pi)^3}{V_i}\Sigma^{(n_1,n_2)(n_1',n_2')}_{(l_1,l_2|m_1,m_2)(l_1',l_2'|m_1',m_2')}(0)\delta^K_{i,j}\,.\tag{D5}$$

We shall now apply this result to the six terms of equation (D1) and write for each of them:

$$\Sigma^{(1,1)(1,1)}_{(1,1|1,1)(1,1|1,1)}(0)=\int d^3r\left[\mathcal{Q}^{(1,1)}_{(1,1|1,1)}(\boldsymbol{r})\right]^2=\int d^3r\left[\mathcal{G}^{(1)}_{(1,1)}(\boldsymbol{r})\right]^4\ ;\tag{D6}$$

$$\Sigma^{(1,1)(2)}_{(1,1|1,1)(1|0)}(0)=\int d^3r\,\mathcal{Q}^{(1,1)}_{(1,1|1,1)}(\boldsymbol{r})\mathcal{Q}^{(2)}_{(1|0)}(\boldsymbol{r})=\int d^3r\left[\mathcal{G}^{(1)}_{(1,1)}(\boldsymbol{r})\right]^2\mathcal{G}^{(2)}_{(1,0)}(\boldsymbol{r})\,;\tag{D7}$$

$$\Sigma^{(1,1)(1)}_{(1,1|1,1)(2|0)}(0)=\int d^3r\,\mathcal{Q}^{(1,1)}_{(1,1|1,1)}(\boldsymbol{r})\mathcal{Q}^{(1)}_{(2|0)}(\boldsymbol{r})=\int d^3r\left[\mathcal{G}^{(1)}_{(1,1)}(\boldsymbol{r})\right]^2\mathcal{G}^{(1)}_{(2,0)}(\boldsymbol{r})\,;\tag{D8}$$

$$\Sigma^{(2)(1)}_{(1|0)(2|0)}(0)=\int d^3r\,\mathcal{Q}^{(2)}_{(1|0)}(\boldsymbol{r})\mathcal{Q}^{(1)}_{(2|0)}(\boldsymbol{r})=\int d^3r\,\mathcal{G}^{(2)}_{(1,0)}(\boldsymbol{r})\mathcal{G}^{(1)}_{(2,0)}(\boldsymbol{r})\,;\tag{D9}$$

$$\Sigma^{(2)(2)}_{(1|0)(1|0)}(0)=\int d^3r\left[\mathcal{Q}^{(2)}_{(1|0)}(\boldsymbol{r})\right]^2=\int d^3r\left[\mathcal{G}^{(2)}_{(1,0)}(\boldsymbol{r})\right]^2\ ;\tag{D10}$$

$$\Sigma^{(1)(1)}_{(2|0)(2|0)}(0)=\int d^3r\left[\mathcal{Q}^{(1)}_{(2|0)}(\boldsymbol{r})\right]^2=\int d^3r\left[\mathcal{G}^{(1)}_{(2,0)}(\boldsymbol{r})\right]^2.\tag{D11}$$

Finally, putting together all these terms we write our final expression for the shell-averaged covariance as

$$\mathrm{Cov}\left[|\tilde{\mathcal{F}}_g(k_i)|^2,|\tilde{\mathcal{F}}_g(k_j)|^2\right]=\frac{2(2\pi)^3}{V_i}\overline{P}^2(k_i)\delta^K_{i,j}\int d^3r\left\{\left[\mathcal{G}^{(1)}_{(1,1)}(\boldsymbol{r})\right]^2+\frac{(1+\alpha)}{\overline{P}(k_i)}\left[\mathcal{G}^{(2)}_{(1,0)}(\boldsymbol{r})+\mathcal{G}^{(1)}_{(2,0)}(\boldsymbol{r})\right]\right\}^2.$$

which is in fact equation (51) from the main text.

APPENDIX E: FUNCTIONAL DERIVATIVES

In order to compute the functional derivatives of $\mathcal{N}$ and $\mathcal{D}$ making up $F[w]$, we must first work out the functional derivatives of the functions $\overline{\mathcal{G}}$ and the normalization A.

E1 Functional derivatives of the $\overline{\mathcal{G}}$ functions and normalization A

For small variations in the path of w, we find that the functional derivative of $\overline{\mathcal{G}}$ can be written:

$$\overline{\mathcal{G}}^{(1)}_{(1,1)}[w+\delta w] = \int \mathrm{d}M \bar{n}(M) b(M) N_{\mathrm{g}}^{(1)}(M) \int \mathrm{d}L \Phi(L|M) \Theta(\boldsymbol{r}|L) \left[w(\boldsymbol{r}, L, M) + \delta w(\boldsymbol{r}, L, M)\right] = \overline{\mathcal{G}}^{(1)}_{(1,1)}[w] + \delta \overline{\mathcal{G}}^{(1)}_{(1,1)}[w];$$

$$\delta \overline{\mathcal{G}}^{(1)}_{(1,1)}[w] \equiv \int \mathrm{d}M \bar{n}(M) b(M) N_{\mathrm{g}}^{(1)}(M) \int \mathrm{d}L \Phi(L|M) \Theta(\boldsymbol{r}|L) \delta w(\boldsymbol{r}, L, M)\,; \tag{E1}$$

$$\overline{\mathcal{G}}^{(2)}_{(1,0)}[w+\delta w] = \int \mathrm{d}M \bar{n}(M) N_{\mathrm{g}}^{(2)}(M) \left\{ \int \mathrm{d}L \Phi(L|M) \Theta(\boldsymbol{r}|L) \left[w(\boldsymbol{r}, L, M) + \delta w(\boldsymbol{r}, L, M)\right] \right\}^2 = \overline{\mathcal{G}}^{(2)}_{(1,0)}[w] + \delta \overline{\mathcal{G}}^{(2)}_{(1,0)}[w];$$

$$\delta \overline{\mathcal{G}}^{(2)}_{(1,0)}[w] \equiv 2 \int \mathrm{d}M \bar{n}(M) N_{\mathrm{g}}^{(2)}(M) \overline{\mathcal{W}}_1(\boldsymbol{r}, M) \int \mathrm{d}L \Phi(L|M) \Theta(\boldsymbol{r}|L) \delta w(\boldsymbol{r}, L, M)\,; \tag{E2}$$

$$\overline{\mathcal{G}}^{(1)}_{(2,0)}[w+\delta w] = \int \mathrm{d}M \bar{n}(M) N_{\mathrm{g}}^{(1)}(M) \int \mathrm{d}L \Phi(L|M) \Theta(\boldsymbol{r}|L) \left[w(\boldsymbol{r}, L, M) + \delta w(\boldsymbol{r}, L, M)\right]^2 = \overline{\mathcal{G}}^{(1)}_{(2,0)}[w] + \delta \overline{\mathcal{G}}^{(1)}_{(2,0)}[w]\,;$$

$$\delta \overline{\mathcal{G}}^{(1)}_{(2,0)}[w] \equiv 2 \int \mathrm{d}M \bar{n}(M) N_{\mathrm{g}}^{(1)}(M) \int \mathrm{d}L \Phi(L|M) \Theta(\boldsymbol{r}|L) w(\boldsymbol{r}, L, M) \delta w(\boldsymbol{r}, L, M). \tag{E3}$$

In the above we have neglected the terms containing $[\delta w]^n$ with $n \geq 2$, and we have used a similar definition to equation (58) and defined:

$$\overline{\mathcal{W}}_l(\boldsymbol{r}, M) = A^{l/2} \mathcal{W}_l(\boldsymbol{r}, M)\,. \tag{E4}$$

Again, for small variations in the value of w, the functional derivative of the normalization constant A can be written:

$$A[w+\delta w] = \int \mathrm{d}^3 r \left(\overline{\mathcal{G}}^{(1)}_{(1,1)}[w+\delta w]\right)^2 = \int \mathrm{d}^3 r \left(\overline{\mathcal{G}}^{(1)}_{(1,1)}[w] + \delta \overline{\mathcal{G}}^{(1)}_{(1,1)}[w]\right)^2 = \int \mathrm{d}^3 r \left[\overline{\mathcal{G}}^{(1)}_{(1,1)}(\boldsymbol{r})\right]^2 + 2 \int \mathrm{d}^3 r\, \overline{\mathcal{G}}^{(1)}_{(1,1)}(\boldsymbol{r}) \delta \overline{\mathcal{G}}^{(1)}_{(1,1)}[w]$$

$$= A[w] + \delta A[w]\,,$$

$$\delta A[w] \equiv 2 \int \mathrm{d}^3 r\, \overline{\mathcal{G}}^{(1)}_{(1,1)}(\boldsymbol{r}) \int \mathrm{d}M\, \bar{n}(M) b(M) N_{\mathrm{g}}^{(1)}(M) \int \mathrm{d}L \Phi(L|M) \Theta(\boldsymbol{r}|L) \delta w(\boldsymbol{r}, L, M)\,. \tag{E5}$$

E2 Functional derivative of $\mathcal{N}[w(\boldsymbol{r}, L, M)]$

Consider equation (56), we may write the functional derivative as

$$\delta \mathcal{N}[w] = 2 \int \mathrm{d}^3 r \left\{ \left[\mathcal{G}^{(1)}_{(1,1)}(\boldsymbol{r})\right]^2 + c \left[\mathcal{G}^{(2)}_{(1,0)}(\boldsymbol{r}) + \mathcal{G}^{(1)}_{(2,0)}(\boldsymbol{r})\right] \right\} \left\{ 2 \mathcal{G}^{(1)}_{(1,1)}(\boldsymbol{r}) \delta \mathcal{G}^{(1)}_{(1,1)}[w] + c \left[\delta \mathcal{G}^{(2)}_{(1,0)}[w] + \delta \mathcal{G}^{(1)}_{(2,0)}[w]\right] \right\}.$$

Using the functional derivatives of equations (E1), (E2), (E3) to calculate the terms in the parenthesis on the right-hand side, we obtain the functional derivative of the numerator $\mathcal{N}$:

$$\delta \mathcal{N}[w] = 4 \int \mathrm{d}^3 r\, \mathrm{d}M \mathrm{d}L \left\{ \left(\left[\mathcal{G}^{(1)}_{(1,1)}(\boldsymbol{r})\right]^2 + c \left[\mathcal{G}^{(2)}_{(1,0)}(\boldsymbol{r}) + \mathcal{G}^{(1)}_{(2,0)}(\boldsymbol{r})\right] \right) \bar{n}(M) N_{\mathrm{g}}^{(1)}(M) \Phi(L|M) \Theta(\boldsymbol{r}|L) \right.$$

$$\left. \times \left[\mathcal{G}^{(1)}_{(1,1)}(\boldsymbol{r}) b(M) + \overline{\mathcal{W}}_1(\boldsymbol{r}, M) N_{\mathrm{g}}^{(2)}(M) / N_{\mathrm{g}}^{(1)}(M) + w(\boldsymbol{r}, L, M) \right] \right\} \delta w(\boldsymbol{r}, L, M). \tag{E6}$$

E3 Functional derivative of $\mathcal{D}[w(\boldsymbol{r}, L, M)]$

Since $\mathcal{D} = A^2$, we have $\delta \mathcal{D}[w] = 2A[w] \delta A[w]$. Using the functional derivative in equation (E5), the functional derivative of $\mathcal{D}[w]$ is given by

$$\delta \mathcal{D}[w] = 4A[w] \int \mathrm{d}^3 r\, \mathrm{d}M \mathrm{d}L \left\{ \mathcal{G}^{(1)}_{(1,1)}(\boldsymbol{r}) \bar{n}(M) b(M) N_{\mathrm{g}}^{(1)}(M) \Phi(L|M) \Theta(\boldsymbol{r}|L) \right\} \delta w(\boldsymbol{r}, L, M)\,. \tag{E7}$$

This paper has been typeset from a TEX/LATEX file prepared by the author.